CRF Errors Corrected by the STIC Systems Branch

Serial Number: 08/206,176

CRF Processing Date: 3/23/94
Edited by: _____
Verified by: _____ (STIC staff)

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically:

  _____
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____.
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically:

  _____
- ☐ Inserted a space between the last nucleic designator and the nucleic number for sequences:

  _____
- ☐ Deleted page numbers in the text of the sequence listing, which is considered invalid text.
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were:

  _____
- ☑ Inserted a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

  Sequence 27
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included:

  _____
- ☐ Deleted extra, invalid, headings used by an applicant, specifically:

  _____
- ☐ Deleted non-ASCII "garbage" at the end of files, and other invalid text, such as a secretary's initials.
- ☐ Inserted mandatory headings, specifically: _____
- ☐ Corrected an obvious error in the response, specifically:

  _____
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically:

  _____
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Other:

  _____

  _____

  _____

***Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.**

8/01/93

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Garner, Ian
Dalrymple, Michael A
Prunkard, Donna E
Foster, Donald C

(ii) TITLE OF INVENTION: Production of Fibrinogen in Transgenic Animals

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ZymoGenetics, Inc.
(B) STREET: 4225 Roosevelt Way, N.E.
(C) CITY: Seattle
(D) STATE: WA
(E) COUNTRY: USA
(F) ZIP: 98105

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Parker, Gary E
(B) REGISTRATION NUMBER: 31-648
(C) REFERENCE/DOCKET NUMBER: 93-15

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 206-547-8080 ext 322
(B) TELEFAX: 206-548-2329

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5943 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: Human Fibrinogen A-alpha chain

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200, 3786..5210)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GTCTAGGAGC CAGCCCCACC CTTAGAAAAG ATG TTT TCC ATG AGG ATC GTC TGC       54
                                 Met Phe Ser Met Arg Ile Val Cys
                                   1               5

CTA GTT CTA AGT GTG GTG GGC ACA GCA TGG GTATGGCCCT TTTCATTTTT         104
Leu Val Leu Ser Val Val Gly Thr Ala Trp
     10                  15

TCTTCTTGCT TTCTCTCTGG TGTTTATTCC ACAAAGAGCC TGGAGGTCAG AGTCTACCTG     164

CTCTATGTCC TGACACACTC TTAGCTTTAT GACCCCAGGC CTGGGAGGAA ATTTCCTGGG     224

TGGGCTTGAC ACCTCAAGAA TACAGGGTAA TATGACACCA AGAGGAAGAT CTTAGATGGA     284

TGAGAGTGTA CAACTACAAG GGAAACTTTA GCATCTGTCA TTCAGTCTTA CCACATTTTG     344

TTTTGTTTTG TTTTAAAAAG GGCAAGAATT ATTTGCCATC CTTGTACCTA TAAAGCCTTG     404

GTGCATTATA ATGCTAGTTA ATGGAATAAA ACATTTTATG GTAAGATTTG TTTTCTTTAG     464

TTATTAATTT CTTGCTACTT GTCCATAATA AGCAGAACTT TTAGTGTTAG TACAGTTTTG     524

CTGAAAGGTT ATTGTTGTGT TTGTCAAGAC AGAAGAAAAA GCAAACGAAT TATCTTTGGA     584

AATATCTTTG CAGTATCAGA AGAGATTAGT TAGTAAGGCA ATACGCTTTT CCGCAGTAAT     644

GGTATTCTTT TAAATTATGA ATCCATCTCT AAAGGTTACA TAGAAACTTG AAGGAGAGAG     704

GAACATTCAG TTAAGATAGT CTAGGTTTTT CTACTGAAGC AGCAATTACA GGAGAAAGAG     764

CTCTACAGTA GTTTTCAACT TTCTGTCTGC AGTCATTAGT AAAAATGAAA AGGTAAAATT     824

TAACTGATTT TATAGATTCA AATAATTTTC CTTTTAGGAT GGATTCTTTA AAACTCCTAA     884

TATTTATCAA ATGCTTATTT AAGTGTCACA CACAGTTAAG AAATTTGTAC ACCTTGTCTC     944
```

```
103  CTTTAATTCT CATAACAACT CCATAAAATG GGTCCTAGGA TTTCCATTTG AAGATAAGAA    1004
104
105  ACCTGAAGCT TGCCGAAGCC CTGTGTCTGC TCTCCTTAAT CTCTGTGAGA GTGCCATCTC    1064
106
107  TTCCTGGGGA CTTGTAGGCA TGCCACTGTC TCCTCTTCTG GCTAACATTG CTGTTGCTCT    1124
108
109  CTTTTGTGTA TGTGAATGAA TCTTTAAAG ACT GCA GAT AGT GGT GAA GGT GAC       1177
110                                 Thr Ala Asp Ser Gly Glu Gly Asp
111                                      20                  25
112
113  TTT CTA GCT GAA GGA GGA GGC GTG CGT GGC CCA AGG GTT GTG GAA AGA       1225
114  Phe Leu Ala Glu Gly Gly Gly Val Arg Gly Pro Arg Val Val Glu Arg
115              30                  35                  40
116
117  CAT CAA TCT GCC TGC AAA GAT TCA GAC TGG CCC TTC TGC TCT GAT GAA       1273
118  His Gln Ser Ala Cys Lys Asp Ser Asp Trp Pro Phe Cys Ser Asp Glu
119           45                  50                  55
120
121  GAC TGG GTAAGCAGTC AGCGGGGGAA GCAGGAGATT CCTTCCCTCT GATGCTAGAG       1329
122  Asp Trp
123       60
124
125  GGGCTCACAG GCTGACCTGA TTGGTCCCAG AAACTTTTTT AAATAGAAAA TAATTGAATA    1389
126
127  GTTACCTACA TAGCAAATAA AGAAAAGGAA CCTACTCCCA AGAGCACTGT TTATTTACCT    1449
128
129  CCCCAACTCT GGATCATTAG TGGGTGAACA GACAGGATTT CAGTTGCATG CTCAGGCAAA    1509
130
131  ACCAGGCTCC TGAGTATTGT GGCCTCAATT TCCTGGCACC TATTTATGGC TAAGTGGACC    1569
132
133  CTCATTCCAG AGTTTCTCTG CGACCTCTAA CTAGTCCTCT TACCTACTTT TAAGCCAACT    1629
134
135  TATCTGGAAG AGAAAGGGTA GGAAGAAATG GGGGCTGCAT GGAAACATGC AAAATTATTC    1689
136
137  TGAATCTGAG AGATAGATCC TTACTGTAAT TTTCTCCCTT CACTTTCAG AAC TAC         1744
138                                                      Asn Tyr
139
140
141  AAA TGC CCT TCT GGC TGC AGG ATG AAA GGG TTG ATT GAT GAA GTC AAT       1792
142  Lys Cys Pro Ser Gly Cys Arg Met Lys Gly Leu Ile Asp Glu Val Asn
143           65                  70                  75
144
145  CAA GAT TTT ACA AAC AGA ATA AAT AAG CTC AAA AAT TCA CTA TTT GAA       1840
146  Gln Asp Phe Thr Asn Arg Ile Asn Lys Leu Lys Asn Ser Leu Phe Glu
147       80                  85                  90
148
149  TAT CAG AAG AAC AAT AAG GAT TCT CAT TCG TTG ACC ACT AAT ATA ATG       1888
150  Tyr Gln Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met
151   95                 100                 105                 110
152
153  GAA ATT TTG AGA GGC GAT TTT TCC TCA GCC AAT  A GTAAGTATTA            1932
```

```
154  Glu Ile Leu Arg Gly Asp Phe Ser Ser Ala Asn
155                  115                 120
156
157  CATATTTACT TCTTTGACTT TATAACAGAA ACAACAAAAA TCCTAAATAA ATATGATATC    1992
158
159  CGCTTATATC TATGACAATT TCATCCCAAA GTACTTAGTG TAGAAACACA TACCTTCATA    2052
160
161  ATATCCCTGA AAATTTTAAG AGGGAGCTTT TGTTTTCGTT ATTTTTTCAA AGTAAAAGAT    2112
162
163  GTTAACTGAG ATTGTTTAAG GTCACAAAAT AAGTCAGAAT TTTGGATTAA AACAAGAATT    2172
164
165  TAAATGTGTT CTTTTCAACA GTATATACTG AAAGTAGGAT GGGTCAGACT CTTTGAGTTG    2232
166
167  ATATTTTTGT TTCTGCTTTG TAAAGGTGAA AACTGAGAGG TCAAGGAACT TGTTCAAAGA    2292
168
169  CACAGAGCTG GGAATTCAAC TCCCAGACTC CACTGAGCTG ATTAGGTAGA TTTTTAAATT    2352
170
171  TAAAATATAG GGTCAAGCTA CGTCATTCTC ACAGTCTACT CATTAGGGTT AGGAAACATT    2412
172
173  GCATTCACTC TGGGCATGGA CAGCGAGTCT AGGGAGTCCT CAGTTTCTCA AGTTTTGCTT    2472
174
175  TGCCTTTTTA CACCTTCACA AACACTTGAC ATTTAAAATC AGTGATGCCA ACACTAGCTG    2532
176
177  GCAAGTGAGT GATCCTGTTG ACCCAAAACA GCTTAGGAAC CATTTCAAAT CTATAGAGTT    2592
178
179  AAAAAGAAAA GCTCATCAGT AAGAAAATCC AATATGTTCA AGTCCCTTGA TTAAGGATGT    2652
180
181  TATAAAATAA TTGAAATGCA ATCAAACCAA CTATTTTAAC TCCAAATTAC ACCTTTAAAA    2712
182
183  TTCCAAAGAA AGTTCTTCTT CTATATTTCT TTGGGATTAC TAATTGCTAT TAGGACATCT    2772
184
185  TAACTGGCAT TCATGGAAGG CTGCAGGGCA TAACATTATC CAAAAGTCAA ATGCCCCATA    2832
186
187  GGTTTTGAAC TCACAGATTA AACTGTAACC AAAATAAAAT TAGGCATATT TACAAGCTAG    2892
188
189  TTTCTTTCTT TCTTTTTTCT CTTTCTTTCT TTCTTTCTTT CTTTCTTTCT TTCTTTCTTT    2952
190
191  CTTTCTTTCT TTCTCCTTCC TTCCTTTCTT CCTTTCTTTT TTGCTGGCAA TTACAGACAA    3012
192
193  ATCACTCAGC AGCTACTTCA ATAACCATAT TTTCGATTTC AG  AC CGT GAT AAT       3065
194                                                   Asn Arg Asp Asn
195                                                              125
196
197  ACC TAC AAC CGA GTG TCA GAG GAT CTG AGA AGC AGA ATT GAA GTC CTG      3113
198  Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu
199                  130                 135                 140
200
201  AAG CGC AAA GTC ATA GAA AAA GTA CAG CAT ATC CAG CTT CTG CAG AAA      3161
202  Lys Arg Lys Val Ile Glu Lys Val Gln His Ile Gln Leu Leu Gln Lys
203              145                 150                 155
204
```

```
205  AAT GTT AGA GCT CAG TTG GTT GAT ATG AAA CGA CTG GAG GTAAGTATGT       3210
206  Asn Val Arg Ala Gln Leu Val Asp Met Lys Arg Leu Glu
207          160                 165                 170
208
209  GGCTGTGGTC CCGAGTGTCC TTGTTTTTGA GTAGAGGGAA AAGGAAGGCG ATAGTTATGC    3270
210
211  ACTGAGTGTC TACTATATGC AGAGAAAAGT GTTATATCCA TCATCTACCT AAAAGTAGGT    3330
212
213  ATTATTTTCC TCACTCCACA GTTGAAGAAA AAAAAATTCA GAGATATTAA GTAAATTTTC    3390
214
215  CAACGTACAT AGATAGTAAT TCAAAGCAAT GTTCAGTCCC TGTCTATTCC AAGCCATTAC    3450
216
217  ATCACCACAC CTCTGAGCCC TCAGCCTGAG TTCACCAAGG ATCATTTAAT TAGCGTTTCC    3510
218
219  TTTGAGAGGG AATAGCACCT TACTCTTGAT CCATTCTGAG GCTAAGATGA ATTAAACAGC    3570
220
221  ATCCATTGCT TATCCTGGCT AGCCCTGCAA TACCCAACAT CTCTTCCACT GAGGGTGCTC    3630
222
223  GATAGGCAGA AAACAGAGAA TATTAAGTGG TAGGTCTCCG AGTCAAAAAA AATGAAACCA    3690
224
225  GTTTCCAGAA GGAAAATTAA CTACCAGGAA CTCAATAGAC GTAGTTTATG TATTTGTATC    3750
226
227  TACATTTTCT CTTTATTTTT CTCCCCTCTC TCTAG GTG GAC ATT GAT ATT AAG       3803
228                                        Val Asp Ile Asp Ile Lys
229                                                        175
230
231  ATC CGA TCT TGT CGA GGG TCA TGC AGT AGG GCT TTA GCT CGT GAA GTA      3851
232  Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val
233              180                 185                 190
234
235  GAT CTG AAG GAC TAT GAA GAT CAG CAG AAG CAA CTT GAA CAG GTC ATT      3899
236  Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile
237          195                 200                 205
238
239  GCC AAA GAC TTA CTT CCC TCT AGA GAT AGG CAA CAC TTA CCA CTG ATA      3947
240  Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile
241      210                 215                 220
242
243  AAA ATG AAA CCA GTT CCA GAC TTG GTT CCC GGA AAT TTT AAG AGC CAG      3995
244  Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln
245  225                 230                 235                 240
246
247  CTT CAG AAG GTA CCC CCA GAG TGG AAG GCA TTA ACA GAC ATG CCG CAG      4043
248  Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln
249                  245                 250                 255
250
251  ATG AGA ATG GAG TTA GAG AGA CCT GGT GGA AAT GAG ATT ACT CGA GGA      4091
252  Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly
253              260                 265                 270
254
255  GGC TCC ACC TCT TAT GGA ACC GGA TCA GAG ACG GAA AGC CCC AGG AAC      4139
```

```
256  Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn
257           275                 280                 285
258
259  CCT AGC AGT GCT GGA AGC TGG AAC TCT GGG AGC TCT GGA CCT GGA AGT      4187
260  Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser
261      290                 295                 300
262
263  ACT GGA AAC CGA AAC CCT GGG AGC TCT GGG ACT GGA GGG ACT GCA ACC      4235
264  Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr
265  305                 310                 315                 320
266
267  TGG AAA CCT GGG AGC TCT GGA CCT GGA AGT GCT GGA AGC TGG AAC TCT      4283
268  Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser
269                 325                 330                 335
270
271  GGG AGC TCT GGA ACT GGA AGT ACT GGA AAC CAA AAC CCT GGG AGC CCT      4331
272  Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro
273             340                 345                 350
274
275  AGA CCT GGT AGT ACC GGA ACC TGG AAT CCT GGC AGC TCT GAA CGC GGA      4379
276  Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly
277           355                 360                 365
278
279  AGT GCT GGG CAC TGG ACC TCT GAG AGC TCT GTA TCT GGT AGT ACT GGA      4427
280  Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly
281      370                 375                 380
282
283  CAA TGG CAC TCT GAA TCT GGA AGT TTT AGG CCA GAT AGC CCA GGC TCT      4475
284  Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser
285  385                 390                 395                 400
286
287  GGG AAC GCG AGG CCT AAC AAC CCA GAC TGG GGC ACA TTT GAA GAG GTG      4523
288  Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val
289                 405                 410                 415
290
291  TCA GGA AAT GTA AGT CCA GGG ACA AGG AGA GAG TAC CAC ACA GAA AAA      4571
292  Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys
293             420                 425                 430
294
295  CTG GTC ACT TCT AAA GGA GAT AAA GAG CTC AGG ACT GGT AAA GAG AAG      4619
296  Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys
297           435                 440                 445
298
299  GTC ACC TCT GGT AGC ACA ACC ACC ACG CGT CGT TCA TGC TCT AAA ACC      4667
300  Val Thr Ser Gly Ser Thr Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr
301      450                 455                 460
302
303  GTT ACT AAG ACT GTT ATT GGT CCT GAT GGT CAC AAA GAA GTT ACC AAA      4715
304  Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys
305  465                 470                 475                 480
306
```

```
307  GAA GTG GTG ACC TCC GAA GAT GGT TCT GAC TGT CCC GAG GCA ATG GAT      4763
308  Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
309                  485                 490                 495
310
311  TTA GGC ACA TTG TCT GGC ATA GGT ACT CTG GAT GGG TTC CGC CAT AGG      4811
312  Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
313              500                 505                 510
314
315  CAC CCT GAT GAA GCT GCC TTC TTC GAC ACT GCC TCA ACT GGA AAA ACA      4859
316  His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
317          515                 520                 525
318
319  TTC CCA GGT TTC TTC TCA CCT ATG TTA GGA GAG TTT GTC AGT GAG ACT      4907
320  Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
321      530                 535                 540
322
323  GAG TCT AGG GGC TCA GAA TCT GGC ATC TTC ACA AAT ACA AAG GAA TCC      4955
324  Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
325  545                 550                 555                 560
326
327  AGT TCT CAT CAC CCT GGG ATA GCT GAA TTC CCT TCC CGT GGT AAA TCT      5003
328  Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
329                  565                 570                 575
330
331  TCA AGT TAC AGC AAA CAA TTT ACT AGT AGC ACG AGT TAC AAC AGA GGA      5051
332  Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
333              580                 585                 590
334
335  GAC TCC ACA TTT GAA AGC AAG AGC TAT AAA ATG GCA GAT GAG GCC GGA      5099
336  Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
337          595                 600                 605
338
339  AGT GAA GCC GAT CAT GAA GGA ACA CAT AGC ACC AAG AGA GGC CAT GCT      5147
340  Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
341      610                 615                 620
342
343  AAA TCT CGC CCT GTC AGA GGT ATC CAC ACT TCT CCT TTG GGG AAG CCT      5195
344  Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
345  625                 630                 635                 640
346
347  TCC CTG TCC CCC TAGACTAAGT TAAATATTTC TGCACAGTGT TCCCATGGCC           5247
348  Ser Leu Ser Pro
349                  645
350
351  CCTTGCATTT CCTTCTTAAC TCTCTGTTAC ACGTCATTGA AACTACACTT TTTTGGTCTG    5307
352
353  TTTTTGTGCT AGACTGTAAG TTCCTTGGGG GCAGGGCCTT TGTCTGTCTC ATCTCTGTAT    5367
354
355  TCCCAAATGC CTAACAGTAC AGAGCCATGA CTCAATAAAT ACATGTTAAA TGGATGAATG    5427
356
357  AATTCCTCTG AAACTCTATT TGAGCTTATT TAGTCAAATT CTTTCACTAT TCAAAGTGTG    5487
```

TGCTATTAGA ATTGTCACCC AACTGATTAA TCACATTTTT AGTATGTGTC TCAGTTGACA 5547

TTTAGGTCAG GCTAAATACA AGTTGTGTTA GTATTAAGTG AGCTTAGCTA CCTGTACTGG 5607

TTACTTGCTA TTAGTTTGTG CAAGTAAAAT TCCAAATACA TTTGAGGAAA ATCCCCTTTG 5667

CAATTTGTAG GTATAAATAA CCGCTTATTT GCATAAGTTC TATCCCACTG TAAGTGCATC 5727

CTTTCCCTAT GGAGGGAAGG AAAGGAGGAA GAAAGAAAGG AAGGGAAAGA AACAGTATTT 5787

GCCTTATTTA ATCTGAGCCG TGCCTATCTT TGTAAAGTTA AATGAGAATA ACTTCTTCCA 5847

ACCAGCTTAA TTTTTTTTTT AGACTGTGAT GATGTCCTCC AAACACATCC TTCAGGTACC 5907

CAAAGTGGCA TTTTCAATAT CAAGCTATCC GGATCC 5943

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 644 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Phe Ser Met Arg Ile Val Cys Leu Val Leu Ser Val Val Gly Thr
 1               5                  10                  15

Ala Trp Thr Ala Asp Ser Gly Glu Gly Asp Phe Leu Ala Glu Gly Gly
            20                  25                  30

Gly Val Arg Gly Pro Arg Val Val Glu Arg His Gln Ser Ala Cys Lys
        35                  40                  45

Asp Ser Asp Trp Pro Phe Cys Ser Asp Glu Asp Trp Asn Tyr Lys Cys
    50                  55                  60

Pro Ser Gly Cys Arg Met Lys Gly Leu Ile Asp Glu Val Asn Gln Asp
65                  70                  75                  80

Phe Thr Asn Arg Ile Asn Lys Leu Lys Asn Ser Leu Phe Glu Tyr Gln
                85                  90                  95

Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met Glu Ile
            100                 105                 110

Leu Arg Gly Asp Phe Ser Ser Ala Asn Asn Arg Asp Asn Thr Tyr Asn
```

```
409          115                 120                 125
410
411   Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu Lys Arg Lys
412       130                 135                 140
413
414   Val Ile Glu Lys Val Gln His Ile Gln Leu Leu Gln Lys Asn Val Arg
415   145                 150                 155                 160
416
417   Ala Gln Leu Val Asp Met Lys Arg Leu Glu Val Asp Ile Asp Ile Lys
418                   165                 170                 175
419
420   Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val
421               180                 185                 190
422
423   Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile
424           195                 200                 205
425
426   Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile
427       210                 215                 220
428
429   Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln
430   225                 230                 235                 240
431
432   Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln
433                   245                 250                 255
434
435   Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly
436               260                 265                 270
437
438   Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn
439           275                 280                 285
440
441   Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser
442       290                 295                 300
443
444   Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr
445   305                 310                 315                 320
446
447   Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser
448                   325                 330                 335
449
450   Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro
451               340                 345                 350
452
453   Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly
454           355                 360                 365
455
456   Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly
457       370                 375                 380
458
459   Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser
```

```
385                 390                 395                 400

Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val
                405                 410                 415

Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys
            420                 425                 430

Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys
        435                 440                 445

Val Thr Ser Gly Ser Thr Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr
    450                 455                 460

Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys
465                 470                 475                 480

Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
                485                 490                 495

Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
            500                 505                 510

His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
        515                 520                 525

Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
    530                 535                 540

Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
545                 550                 555                 560

Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
                565                 570                 575

Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
            580                 585                 590

Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
        595                 600                 605

Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
    610                 615                 620

Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
625                 630                 635                 640

Ser Leu Ser Pro


(2) INFORMATION FOR SEQ ID NO:3:
```

```
511
512         (i) SEQUENCE CHARACTERISTICS:
513               (A) LENGTH: 8878 base pairs
514               (B) TYPE: nucleic acid
515               (C) STRANDEDNESS: double
516               (D) TOPOLOGY: linear
517
518        (ii) MOLECULE TYPE: DNA (genomic)
519
520
521       (vii) IMMEDIATE SOURCE:
522               (B) CLONE: human fibrinogen B-beta chain
523
524        (ix) FEATURE:
525               (A) NAME/KEY: misc_RNA
526               (B) LOCATION: 1..469
527
528        (ix) FEATURE:
529               (A) NAME/KEY: exon
530               (B) LOCATION: 470..583
531
532        (ix) FEATURE:
533               (A) NAME/KEY: intron
534               (B) LOCATION: 584..3257
535
536        (ix) FEATURE:
537               (A) NAME/KEY: exon
538               (B) LOCATION: 3258..3449
539
540        (ix) FEATURE:
541               (A) NAME/KEY: intron
542               (B) LOCATION: 3450..3938
543
544        (ix) FEATURE:
545               (A) NAME/KEY: exon
546               (B) LOCATION: 3939..4122
547
548        (ix) FEATURE:
549               (A) NAME/KEY: intron
550               (B) LOCATION: 4123..5042
551
552        (ix) FEATURE:
553               (A) NAME/KEY: exon
554               (B) LOCATION: 5043..5270
555
556        (ix) FEATURE:
557               (A) NAME/KEY: intron
558               (B) LOCATION: 5271..5830
559
560        (ix) FEATURE:
561               (A) NAME/KEY: exon
```

562 (B) LOCATION: 5831..5944
563
564 (ix) FEATURE:
565 (A) NAME/KEY: intron
566 (B) LOCATION: 5945..6632
567
568 (ix) FEATURE:
569 (A) NAME/KEY: exon
570 (B) LOCATION: 6633..6758
571
572 (ix) FEATURE:
573 (A) NAME/KEY: intron
574 (B) LOCATION: 6759..6966
575
576 (ix) FEATURE:
577 (A) NAME/KEY: exon
578 (B) LOCATION: 6967..7252
579
580 (ix) FEATURE:
581 (A) NAME/KEY: intron
582 (B) LOCATION: 7253..7870
583
584
585 (ix) FEATURE:
586 (A) NAME/KEY: exon
587 (B) LOCATION: 7871..8102
588
589 (ix) FEATURE:
590 (A) NAME/KEY: 3'UTR
591 (B) LOCATION: 8103..8537
592
593 (ix) FEATURE:
594 (A) NAME/KEY: misc_RNA
595 (B) LOCATION: 8538..8878
596
597 (ix) FEATURE:
598 (A) NAME/KEY: CDS
599 (B) LOCATION: join(470..583, 3258..3449, 3939..4122, 5043..5270,
600 5831..5944, 6633..6758, 6967..7252, 7871..8102)
601
602
603 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
604
605 GAATTCATGC CCCTTTTGAA ATAGACTTAT GTCATTGTCA GAAAACATAA GCATTTATGG 60
606
607 TATATCATTA ATGAGTCACG ATTTTAGTGG TTGCCTTGTG AGTAGGTCAA ATTTACTAAG 120
608
609 CTTAGATTTG TTTTCTCACA TATTCTTTCG GAGCTTGTGT AGTTTCCACA TTAATTTACC 180
610
611 AGAAACAAGA TACACACTCT CTTTGAGGAG TGCCCTAACT TCCCATCATT TTGTCCAATT 240
612

```
613  AAATGAATTG AAGAAATTTA ATGTTTCTAA ACTAGACCAA CAAAGAATAA TAGTTGTATG     300
614
615  ACAAGTAAAT AAGCTTTGCT GGGAAGATGT TGCTTAAATG ATAAAATGGT TCAGCCAACA     360
616
617  AGTGAACCAA AAATTAAATA TTAACTAAGG AAAGGTAACC ATTTCTGAAG TCATTCCTAG     420
618
619  CAGAGGACTC AGATATATAT AGGATTGAAG ATCTCTCAGT TAAGTCTAC ATG AAA          475
620                                                        Met Lys
621                                                          1
622
623  AGG ATG GTT TCT TGG AGC TTC CAC AAA CTT AAA ACC ATG AAA CAT CTA        523
624  Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys His Leu
625            5                  10                  15
626
627  TTA TTG CTA CTA TTG TGT GTT TTT CTA GTT AAG TCC CAA GGT GTC AAC        571
628  Leu Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Gln Gly Val Asn
629       20                  25                  30
630
631  GAC AAT GAG GAG GTGAATTTTT TAAAGCATTA TTATATTATT AGTAGTATTA            623
632  Asp Asn Glu Glu
633   35
634
635  TTAATATAAG ATGTAACATA ATCATATTAT GTGCTTATTT TAATGAAATT AGCATTGCTT     683
636
637  ATAGTTATGA AATGGAATTG TTAACCTCTG ACTTATTGTA TTTAAAGAAT GTTTCATAGT     743
638
639  ATTTCTTATA TAAAAACAAA GTAATTTCTT GTTTTCTAGT TTATCACCTT TGTTTTCTTA     803
640
641  AGATGAGGAT GGCTTAGCTA ATGTAAGATG TGTTTTTCTC ACTTGCTATT CTGAGTACTG     863
642
643  TGATTTTCAT TTACTTCTAG CAATACAGGA TTACAATTAA GAGGACAAGA TCTGAAAATC     923
644
645  TCACAAACTA TAAAATAATA AAAGAGCAGA ATTTTAAGAT AAAAGAAACT GGTGGTAGGT     983
646
647  AGATTGTTCT TTGGTGAAGG AAGGTAATAT ATATTGTTAC TGAGATTACT ATTTATAAAA    1043
648
649  ATTATAACTA AGCCTAAAAG CAAAATACAT CAAGTGTAAT GATAGAAAAT GAAATATTGC    1103
650
651  TTTTTTCAGA TGAAAAGTTC AAATTAGAGT TAGTGTGTAT TGTTATTATT AATAGTTATG    1163
652
653  AAACACGGTT CAGTCTAATT TATTTATTTG TAGAACAGTT TGTCCTCAAC TATTATTTTT    1223
654
655  GCTGACTTAT TGCTGTTAAT TTGCAGTTAC TAAAAATACA GAAATGCATT TAGGACAATG    1283
656
657  GATATTTAAG AAATTTAAAT TTTATCATCA AACGTATCAT GGCCAAATTT CTTACATATA    1343
658
659  GCATAGTATC ATTAAACTAG AAATAAGAAT ACACAATAAT ATTTAAATGA AGTGATTCAT    1403
660
661  TTCGGATCAT TATTGAGTTT CAAGGGAACT TGAGTGTTGT ACTTATCAGA CTCTACATGT    1463
662
663  AAGAACATAT AGTTAATCTG GTTGTGTGTG TAAAAACATA TGGTTAATCT GGTTAAGTCT    1523
```

664
665 GGTTAATCAT ATTAGGTAAG AAAAATGTAA AGAATGTGTA AGACGAAATT TTTGTAAAGT 1583
666
667 ACTCTGCAAA GCACTTTCAC ATTTCTGCTT ATCAACTAAA CCTCACAGAG ATAGTTTAAT 1643
668
669 AGTTTAGGCT TTAAAATGGA TTTTGATTAT TCAACAAGTG GCCTTCATAA TTTCTTTAAG 1703
670
671 TGTTTTTCTT TAAGTATATA CTTTCTTTAA ATATTTTTTA AAATTTCCTT TTCTCTAGTA 1763
672
673 AAGCCAGACC ATCCATGCTA CCTCTCTAGT GGCACTCTGA AATAAAAAGA AAATAGTTTT 1823
674
675 CTCTGTTATA ATTGTATTTG TAATAAGCAG ATGAATCACA TTTCTTAAAA TTTGTTTTAG 1883
676
677 AGAGGGTAAG CTCTGACTAG GACCATGACT TCAATGTGAA ATATGTATAT ATCCTCCGAA 1943
678
679 TCTTTACATA TTAAGAATGT ATATAGTCAA CTGGTTAAAC AGGAAAATCT GGAACAGCCT 2003
680
681 GGCTGGGTTT TAATCTTAGC ACCATCCTAC TAAATGTTAA ATAATATTAT AATCTAATGA 2063
682
683 ATAAATGACA ATGCAATTCC AAATAGAGTT CATCTGATGA CTTCTAGACT CACAAAATTG 2123
684
685 CAAGAGAGCT CAGTTGTTGC TCAGTTGTTC CAAATCATGT CGTTTGTTAA TTTGTAATTA 2183
686
687 AGCTCCAAAG GATGTATAGC TACTGACAAA AAAAAAAATG AGAATGTAGT TAATCCAAAT 2243
688
689 CAAAACTTTC CTATTGCAAT GCGTATTTTC TGCTTCATTA TCCTTTAATA TAATATTTTA 2303
690
691 AGTTAGCAAG TAATTTTAAT TACAATGCAC AAGCCTTGAG AATTATTTTA AATATAAGAA 2363
692
693 AATCATAATG TTTGATAAAG AAATCATGTA AGAAATTTCA AGATAATGGT TTAACAAATA 2423
694
695 ATTTTGTTGA TAGAAGATAA GACTAAAAGT GAAATTCGAA GTGGAGAGGA CACTTAAACT 2483
696
697 GTAGTACTTG TTATGTGTGA TTCCAGTAAA AATAGTAATG AGCACTTATT ATTGCCAAGT 2543
698
699 ACTGTTCTGA GGGTACCATA TGCAATAAGT TATTTAATCC TTACAATAAT CTTGTAAGGC 2603
700
701 AGATTCAAAC TATCATTACA CTTATTTTAC AGATGAGAAA ACTGGGGCAC AGATAAAGCA 2663
702
703 ACTTGCCCAA GGTCTCATAG CTGTAAGTCA ACCCTACGGT CAAGACCTAC AAGTAGCCGA 2723
704
705 GCTCCAGAGT ACATTATGAG GGTCAAAGAT TGTCTTATTA CAAATAAATT CCAAGTAGAA 2783
706
707 TCAACCTTTA ATAAGTCTTT AATGTCTCTT AAATATGTTT ATATAGGAGT CTAATCACCA 2843
708
709 ATTCACAAAA ATGAAAGTAG GGAAATGATT AACAATAATC ATAGGAATCT AACAATCCAA 2903
710
711 GTGGCTTGAG AATATTCATT CTTCTTGACA GTATAGATTC TTTACAATTT CGTAAGTTCC 2963
712
713 AATGTATGTT TTAGGAATAT GAGGTCATTA CTATTCATAA TCTGATACAG CTTTATCCTA 3023
714

```
715  AGGCCTCTCT TTAAAAACTA CACTGCATCA TAGCTTTTTT GTGCAGTTGG TCTTTCTACT    3083
716
717  GTTACTGAAC AGTAAGCAAC CTACAGATTC ACTATCACCA ACCAGCCAGT TGATGGATCT    3143
718
719  TAAGCAAATT ATCAAGCTTG TGATAACCTA AATTATAAAA TGAGGGTGTT GGAATAGTTA    3203
720
721  CATTCCAAAT CTTCTATAAC ACTCTGTATT ATATTTCTGC CTCATTCCTT GTAG GGT      3260
722                                                              Gly
723
724
725  TTC TTC AGT GCC CGT GGT CAT CGA CCC CTT GAC AAG AAG AGA GAA GAG      3308
726  Phe Phe Ser Ala Arg Gly His Arg Pro Leu Asp Lys Lys Arg Glu Glu
727   40                  45                  50                  55
728
729  GCT CCC AGC CTG AGG CCT GCC CCA CCG CCC ATC AGT GGA GGT GGC TAT      3356
730  Ala Pro Ser Leu Arg Pro Ala Pro Pro Pro Ile Ser Gly Gly Gly Tyr
731                  60                  65                  70
732
733  CGG GCT CGT CCA GCC AAA GCA GCT GCC ACT CAA AAG AAA GTA GAA AGA      3404
734  Arg Ala Arg Pro Ala Lys Ala Ala Ala Thr Gln Lys Lys Val Glu Arg
735              75                  80                  85
736
737  AAA GCC CCT GAT GCT GGA GGC TGT CTT CAC GCT GAC CCA GAC CTG          3449
738  Lys Ala Pro Asp Ala Gly Gly Cys Leu His Ala Asp Pro Asp Leu
739           90                  95                  100
740
741  GTGGGTGCAC TGATGTTTCT TGCAGTGGTG GCTCTCTCAT GCAGAGAAAG CCTGTAGTCA    3509
742
743  TGGCAGTCTG CTAATGTTTC ACTGACCCAC ATTACCATCA CTGTTATTTT GTTTGTTTAT    3569
744
745  TTTGGAAATA AAATTCAAAA CATAAACATA TTGGGCCTTT GGTTTAGGCT TTCTTTCTTG    3629
746
747  TTTTCTTTGG TCTGGGCCCA AAATTTCAAA TTAGGATATG TGGGTGCCAC CTTTCCATTT    3689
748
749  GTATTTTGCC ACTGCCTTTG TTTAGTTGGT AAAATTTTCA TAGCCCAATT ATATTTTTTC    3749
750
751  TGGGGTAAGT AATATTTTAA ATCTCTATGA GAGTATGATG ATGACTTTCG AATTTCTGGT    3809
752
753  CTTACAGAAA ACCAAATAAT AAATTTTTAT GTTGGCTAAT CGTATCGCTG AATTTTCCTA    3869
754
755  TGTGCTATTT TAACAAATGT CCATGACCCA AATCCTTCAT CTAATGCCTG CTATTTTCTT    3929
756
757  TGTTTTTAG GGG GTG TTG TGT CCT ACA GGA TGT CAG TTG CAA GAG GCT         3977
758            Gly Val Leu Cys Pro Thr Gly Cys Gln Leu Gln Glu Ala
759                  105                 110                 115
760
761  TTG CTA CAA CAG GAA AGG CCA ATC AGA AAT AGT GTT GAT GAG TTA AAT      4025
762  Leu Leu Gln Gln Glu Arg Pro Ile Arg Asn Ser Val Asp Glu Leu Asn
763                  120                 125                 130
764
765  AAC AAT GTG GAA GCT GTT TCC CAG ACC TCC TCT TCT TCC TTT CAG TAC      4073
```

```
766  Asn Asn Val Glu Ala Val Ser Gln Thr Ser Ser Ser Ser Phe Gln Tyr
767              135                 140                 145
768
769  ATG TAT TTG CTG AAA GAC CTG TGG CAA AAG AGG CAG AAG CAA GTA AAA  G   4122
770  Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys Gln Val Lys
771          150                 155                 160
772
773  GTAGATATCC TTGTGCTTTC CATTCGATTT TCAGCTATAA AATTGGAACC GTTAGACTGC    4182
774
775  CACGAGAATG CATGGTTGTG AGAAGATTAA CATTTCTGGG TTAGTGAATA GCATTCATAC    4242
776
777  GCTTTTGGGC ACCTTCCCCT GCAACTTGCC AGATAAGCAC TATTCAGCTC TTATTCCCAG    4302
778
779  TCTGACATCA GCAAGTGTGA TTTTCTATGA AAAATTCTAC TATGACTCCT TATTTTAAGT    4362
780
781  ATACAAGAAA CTTGTGACTC AGAAGATAAT ATTTACAGAG TGGAAAAAAA CCCCTAGCAT    4422
782
783  TTATAGTTTT AACATTTGAG GTTTTGAATG AGAGAGTTAT CCATAATATA TTCAATTGTG    4482
784
785  TTGTGGATAA TGACACCTAA CCTGTGAATC TTGAGGTCAG AATGTTGAGT GCTGTTGACT    4542
786
787  TGGTGGTCAG GAAACAGCTA GTGCGTGAGC CTGGCACAGG CATCTCAGTG AGTAGCATAC    4602
788
789  CCACAGTTGG AAATTTTTCA AAGAAATCAA AGGAATCATG ACATCTTATA AATTTCAAGG    4662
790
791  TTCTGCTATA CTTATGTGAA ATGGATAAAT AAATCAAGCA TATCCACTCT GTAAGATTGA    4722
792
793  ACTTCTCAGA TGGAAGACCC CAATACTGCT TTCTCCTCTT TTCCCTCACC AAAGAAATAA    4782
794
795  ACAACCTATT TCATTTATTA CTGGACACAA TCTTTAGCGT ATACCTATGG TAAATTACTA    4842
796
797  GTATGGTGGT TAGGATTTAT GTTAATTTGT ATATGTCATG CGCCAAATCA TTTCCACTAA    4902
798
799  ATATGACTAT ATATCATAAC TGCTTGGTGA TAGCTCAGTG TTTAATAGTT TATTCTCAGA    4962
800
801  AAATCAAAAT TGTATAGTTA AATACATTAG TTTTATGAGG CAAAAATGCT AACTATTTCT    5022
802
803  ACATAATTTC ATTTTTCCAG  AT AAT GAA AAT GTA GTC AAT GAG TAC TCC         5071
804                       Asp Asn Glu Asn Val Val Asn Glu Tyr Ser
805                           165                 170
806
807  TCA GAA CTG GAA AAG CAC CAA TTA TAT ATA GAT GAG ACT GTG AAT AGC      5119
808  Ser Glu Leu Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser
809      175                 180                 185
810
811  AAT ATC CCA ACT AAC CTT CGT GTG CTT CGT TCA ATC CTG GAA AAC CTG      5167
812  Asn Ile Pro Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu
813  190                 195                 200                 205
814
815  AGA AGC AAA ATA CAA AAG TTA GAA TCT GAT GTC TCA GCT CAA ATG GAA      5215
816  Arg Ser Lys Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu
```

```
817                 210                 215                 220
818
819  TAT TGT CGC ACC CCA TGC ACT GTC AGT TGC AAT ATT CCT GTG GTG TCT      5263
820  Tyr Cys Arg Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser
821              225                 230                 235
822
823  GGC AAA  G GTAACTGATT CATAAACATA TTTTTAGAGA GTTCCAGAAG AACTCACACA    5320
824  Gly Lys
825
826
827  CCAAAAATAA GAGAACAACA ACAACAACAA AAATGCTAAG TGGATTTTCC CAACAGATCA    5380
828
829  TAATGACATT ACAGTACATC ATAAAAATAT CCTTAGCCAG TTGTGTTTTG GACTGGCCTG    5440
830
831  GTGCATTTGC TGGTTTTGAT GAGCAGGATG GGGCACAGGT AGTCCCAGGG GTGGCTGATG    5500
832
833  TGTGCATCTG CGTACTGGCT TGAACAGATG GCAGAACCAC AGATAGATGT AGAAGTTTCT    5560
834
835  CCATTTTGTG TGTTCTGGGA GCTCATGGAT ATTCCAGGAC ACAAAAGGTG GAGAAGAGCT    5620
836
837  TTGTTCATCC TCTTAGCAGA TAAACGTCCT CAAAACTGGG TTGGACTTAC TAAAGTAAAA    5680
838
839  TGAAAATCTA ATATTTGTTA TATTATTTTC AAAGGTCTAT AATAACACAC TCCTTAGTAA    5740
840
841  CTTATGTAAT GTTATTTTAA AGAATTGGTG ACTAAATACA AAGTAATTAT GTCATAAACC    5800
842
843  CCTGAACATA ATGTTGTCTT ACATTTGCAG  AA TGT GAG GAA ATT ATC AGG AAA      5853
844                                   Glu Cys Glu Glu Ile Ile Arg Lys
845                                   240                 245
846
847  GGA GGT GAA ACA TCT GAA ATG TAT CTC ATT CAA CCT GAC AGT TCT GTC      5901
848  Gly Gly Glu Thr Ser Glu Met Tyr Leu Ile Gln Pro Asp Ser Ser Val
849          250                 255                 260
850
851  AAA CCG TAT AGA GTA TAC TGT GAC ATG AAT ACA GAA AAT GGA  G           5944
852  Lys Pro Tyr Arg Val Tyr Cys Asp Met Asn Thr Glu Asn Gly
853      265                 270                 275
854
855  GTAAGCTTTC GACAGTTGTT GACCTGTTGA TCTGTAATTA TTTGGATACC GTAAAATGCC    6004
856
857  AGGAAACAAG GCCAGGTGTG GTGGCTCATA CCTGTAATTC CAGCACCTTG GGAGGCCAAA    6064
858
859  GTGGGCTGAT AGCTTGAGCC TAGGAGTTTG AAACTAGCCT GGGCAACATA ATGAGACCCT    6124
860
861  AACTCTACAA AAAAAAAAAA AATACCAAAA AAAAAAAAAA AATCAGCTGT GTTGGTAGTA    6184
862
863  TGTGCCTGTA GTCCCAGCTA TCCAGGAGGC TGAGATGGGA GATCACCTGA GCCCACAACC    6244
864
865  TGGAGTCTTG ATCATGCTAC TGAACTGTAG CCTGGGCAAC AGAGGATAGT GAGATCCTGT    6304
866
867  CTCAAAAAAA AAAATTAATT AAAAAGCCAG GAAACAAGAC TTAGCTCTAA CATCTAACAT    6364
```

```
868
869  AGCTGACAAA GGAGTAATTT GATGTGGAAT TCAACCTGAT ATTTAAAAGT TATAAAATAT    6424
870
871  CTATAATTCA CAATTTGGGG TAAGATAAAG CACTTGCAGT TTCCAAAGAT TTTACAAGTT    6484
872
873  TACCTCTCAT ATTTATTTCC TTATTGTGTC TATTTTAGAG CACCAAATAT ATACTAAATG    6544
874
875  GAATGGACAG GGGATTCAGA TATTATTTTC AAAGTGACAT TATTTGCTGT TGGTTAATAT    6604
876
877  ATGCTCTTTT TGTTTCTGTC AACCAAAG  GA TGG ACA GTG ATT CAG AAC CGT       6655
878                                 Gly Trp Thr Val Ile Gln Asn Arg
879                                          280                 285
880
881  CAA GAC GGT AGT GTT GAC TTT GGC AGG AAA TGG GAT CCA TAT AAA CAG      6703
882  Gln Asp Gly Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln
883                  290                 295                 300
884
885  GGA TTT GGA AAT GTT GCA ACC AAC ACA GAT GGG AAG AAT TAC TGT GGC      6751
886  Gly Phe Gly Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly
887              305                 310                 315
888
889  CTA CCA  G GTAACGAACA GGCATGCAAA ATAAAATCAT TCTATTTGAA ATGGGATTTT    6808
890  Leu Pro
891
892
893  TTTTAATTAA AAAACATTCA TTGTTGGAAG CCTGTTTTAG GCAGTTAAGA GGAGTTTCCT    6868
894
895  GACAAAAATG TGGAAGCTAA AGATAAGGGA AGAAAGGCAG TTTTTAGTTT CCCAAAATTT    6928
896
897  TATTTTTGGT GAGAGATTTT ATTTTGTTTT TCTTTTAG  GT GAA TAT TGG CTT        6980
898                                           Gly Glu Tyr Trp Leu
899                                           320
900
901  GGA AAT GAT AAA ATT AGC CAG CTT ACC AGG ATG GGA CCC ACA GAA CTT      7028
902  Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly Pro Thr Glu Leu
903  325                 330                 335                 340
904
905  TTG ATA GAA ATG GAG GAC TGG AAA GGA GAC AAA GTA AAG GCT CAC TAT      7076
906  Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val Lys Ala His Tyr
907                  345                 350                 355
908
909  GGA GGA TTC ACT GTA CAG AAT GAA GCC AAC AAA TAC CAG ATC TCA GTG      7124
910  Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr Gln Ile Ser Val
911              360                 365                 370
912
913  AAC AAA TAC AGA GGA ACA GCC GGT AAT GCC CTC ATG GAT GGA GCA TCT      7172
914  Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met Asp Gly Ala Ser
915          375                 380                 385
916
917  CAG CTG ATG GGA GAA AAC AGG ACC ATG ACC ATT CAC AAC GGC ATG TTC      7220
918  Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His Asn Gly Met Phe
```

```
919      390                 395                 400
920
921  TTC AGC ACG TAT GAC AGA GAC AAT GAC GGC  TG  GTATGTGTGG              7262
922  Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly  Trp
923  405                 410                 415
924
925  CACTCTTTGC TCCTGCTTTA AAAATCACAC TAATATCATT ACTCAGAATC ATTAACAATA    7322
926
927  TTTTTAATAG CTACCACTTC CTGGGCACTT ACTGTCAGCC ACTGTCCTAA GCTCTTTATG    7382
928
929  CATCACTCGA AAGCATTTCA ACTATAAGGT AGACATTCTT ATTCTCATTT TACAGATGAG    7442
930
931  ATTTAGAGAG ATTACGTGAT TTGTCCAATG TCACACAACT ACCCAGAGAT AAAACTAGAA    7502
932
933  TTTGAGCACA GTTACTTTCT GAATAATGAG CATTTAGATA AATACCTATA TCTCTATATT    7562
934
935  CTAAAGTGTG TGTGAAAACT TTCATTTTCA TTTCCAGGGT TCTCTGATAC TAAGGGTTGT    7622
936
937  AAAAGCTATT ATTCCAGTAT AAAGTAACAA ACACAGTCCC TAGATGGATT GCCACAAAGG    7682
938
939  CCCAGTTATC TCTCTTTCTT GCTATAGGGC ACAGGAGGTC TTTGGTGTAT TAGTGTGACT    7742
940
941  CTATGTATAG CACCCAAAGG AAAGACTACT GTGCACACGA GTGTAGCAGT CTTTTATGGG    7802
942
943  TAATCTGCAA AACGTAACTT GACCACCGTA GTTCTGTTTC TAATAACGCC AAACACATTT    7862
944
945  TCTTTCAG G TTA ACA TCA GAT CCC AGA AAA CAG TGT TCT AAA GAA GAC       7910
946              Leu Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp
947                              420                 425
948
949  GGT GGT GGA TGG TGG TAT AAT AGA TGT CAT GCA GCC AAT CCA AAC GGC      7958
950  Gly Gly Gly Trp Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly
951      430                 435                 440
952
953  AGA TAC TAC TGG GGT GGA CAG TAC ACC TGG GAC ATG GCA AAG CAT GGC      8006
954  Arg Tyr Tyr Trp Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly
955  445                 450                 455                 460
956
957  ACA GAT GAT GGT GTA GTA TGG ATG AAT TGG AAG GGG TCA TGG TAC TCA      8054
958  Thr Asp Asp Gly Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser
959                  465                 470                 475
960
961  ATG AGG AAG ATG AGT ATG AAG ATC AGG CCC TTC TTC CCA CAG CAA TAGTCCCCAA   8109
962  Met Arg Lys Met Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln
963              480                 485                 490
964
965  TACGTAGATT TTTGCTCTTC TGTATGTGAC AACATTTTTG TACATTATGT TATTGGAATT    8169
966
967  TTCTTTCATA CATTATATTC CTCTAAAACT CTCAAGCAGA CGTGAGTGTG ACTTTTTGAA    8229
968
969  AAAAGTATAG GATAAATTAC ATTAAAATAG CACATGATTT TCTTTTGTTT TCTTCATTTC    8289
```

TCTTGCTCAC CCAAGAAGTA ACAAAAGTAT AGTTTTGACA GAGTTGGTGT TCATAATTTC 8349

AGTTCTAGTT GATTGCGAGA ATTTTCAAAT AAGGAAGAGG GGTCTTTTAT CCTTGTCGTA 8409

GGAAAACCAT GACGGAAAGG AAAAACTGAT GTTTAAAAGT CCACTTTTAA AACTATATTT 8469

ATTTATGTAG GATCTGTCAA AGAAAACTTC CAAAAAGATT TATTAATTAA ACCAGACTCT 8529

GTTGCAATAA GTTAATGTTT TCTTGTTTTG TAATCCACAC ATTCAATGAG TTAGGCTTTG 8589

CACTTGTAAG GAAGGAGAAG CGTTCACAAC CTCAAATAGC TAATAAACCG GTCTTGAATA 8649

TTTGAAGATT TAAAATCTGA CTCTAGGACG GGCACGGTGG CTCACGACTA TAATCCCAAC 8709

ACTTTGGGAG GCTGAGGCGG GCGGTCACAA GGTCAGGAGT TCAAGACCAG CCTGACCAAT 8769

ATGGTGAAAC CCCATCTCTA CTAAAAATAC AAAAATTAGC CAGGCGTGGT GGCAGGTGCC 8829

TGTAGGTCCC AGCTAGCCTG TGAGGTGGAG ATTGCATTGA GCCAAGATC 8878

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 491 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Lys Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys
  1               5                  10                  15

His Leu Leu Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Gln Gly
             20                  25                  30

Val Asn Asp Asn Glu Glu Gly Phe Phe Ser Ala Arg Gly His Arg Pro
         35                  40                  45

Leu Asp Lys Lys Arg Glu Glu Ala Pro Ser Leu Arg Pro Ala Pro Pro
     50                  55                  60

Pro Ile Ser Gly Gly Gly Tyr Arg Ala Arg Pro Ala Lys Ala Ala Ala
 65                  70                  75                  80

Thr Gln Lys Lys Val Glu Arg Lys Ala Pro Asp Ala Gly Gly Cys Leu
                 85                  90                  95
```

```
1021  His Ala Asp Pro Asp Leu Gly Val Leu Cys Pro Thr Gly Cys Gln Leu
1022               100                 105                 110
1023
1024  Gln Glu Ala Leu Leu Gln Gln Glu Arg Pro Ile Arg Asn Ser Val Asp
1025          115                 120                 125
1026
1027  Glu Leu Asn Asn Asn Val Glu Ala Val Ser Gln Thr Ser Ser Ser Ser
1028      130                 135                 140
1029
1030  Phe Gln Tyr Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys
1031  145                 150                 155                 160
1032
1033
1034  Gln Val Lys Asp Asn Glu Asn Val Val Asn Glu Tyr Ser Ser Glu Leu
1035                  165                 170                 175
1036
1037  Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser Asn Ile Pro
1038               180                 185                 190
1039
1040  Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu Arg Ser Lys
1041          195                 200                 205
1042
1043  Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu Tyr Cys Arg
1044      210                 215                 220
1045
1046  Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser Gly Lys Glu
1047  225                 230                 235                 240
1048
1049  Cys Glu Glu Ile Ile Arg Lys Gly Gly Glu Thr Ser Glu Met Tyr Leu
1050                  245                 250                 255
1051
1052  Ile Gln Pro Asp Ser Ser Val Lys Pro Tyr Arg Val Tyr Cys Asp Met
1053               260                 265                 270
1054
1055  Asn Thr Glu Asn Gly Gly Trp Thr Val Ile Gln Asn Arg Gln Asp Gly
1056          275                 280                 285
1057
1058  Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln Gly Phe Gly
1059      290                 295                 300
1060
1061  Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly Leu Pro Gly
1062  305                 310                 315                 320
1063
1064  Glu Tyr Trp Leu Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly
1065                  325                 330                 335
1066
1067  Pro Thr Glu Leu Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val
1068               340                 345                 350
1069
1070  Lys Ala His Tyr Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr
1071          355                 360                 365
```

```
1072
1073  Gln Ile Ser Val Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met
1074      370                 375                 380
1075
1076  Asp Gly Ala Ser Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His
1077  385                 390                 395                 400
1078
1079  Asn Gly Met Phe Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly Trp Leu
1080                  405                 410                 415
1081
1082  Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp Gly Gly Gly Trp
1083              420                 425                 430
1084
1085  Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly Arg Tyr Tyr Trp
1086          435                 440                 445
1087
1088  Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly Thr Asp Asp Gly
1089      450                 455                 460
1090
1091  Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser Met Arg Lys Met
1092  465                 470                 475                 480
1093
1094  Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln
1095                  485                 490
1096
1097
1098  (2) INFORMATION FOR SEQ ID NO:5:
1099
1100       (i) SEQUENCE CHARACTERISTICS:
1101            (A) LENGTH: 10564 base pairs
1102            (B) TYPE: nucleic acid
1103            (C) STRANDEDNESS: double
1104            (D) TOPOLOGY: linear
1105
1106      (ii) MOLECULE TYPE: DNA (genomic)
1107
1108
1109     (vii) IMMEDIATE SOURCE:
1110            (B) CLONE: human fibrinogen gamma chain
1111
1112      (ix) FEATURE:
1113            (A) NAME/KEY: CDS
1114            (B) LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
1115                  ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
1116                  ..7703, 9342..9571)
1117
1118
1119      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
1120
1121  CTACACACTT CTTGAAGGCA AAGGCAATGC TGAAGTCACC TTTCATGTTC AAATCATATT      60
1122
```

1123 AAAAAGTTAG CAAGATGTAA TTATCAGTGT ACTATGTAAA TCTTTGTGAA TGATCAATAA 120
1124
1125 TTACATATTT TCATTATATA TATTTTAGTA GATAATATTT ATATACATTC AACATTCTAA 180
1126
1127 ATATAGAAAG TTTACAGAGA AAAATAAAGC CTTTTTTTCC AATCCTGTCC TCCACCTCTG 240
1128
1129 CATCCCATTC TTCTTCACAG AGGCAACTGA TTCAAGTCAT TACATAGTTA TTGAGTGTTA 300
1130
1131 ACTACAACTA TGTTAAGTAC AGCTATATAT GTTAGATGCC GTAGCCACAG AAATCAGTTT 360
1132
1133 ACAATCTAAT GCAGTGGATA CAGCATGTAT ACATATAATA TAAGGTTGCT ACAAATGCTA 420
1134
1135 TCTGAGGTAG AGCTGTTTGA AAGAATACTA ATACTTAAAT GTTTAATTCA ACTGACTTGA 480
1136
1137 TTGACAACTG ATTAGCTGAG TGGAAAAGAT GGATGAGAAA GATTGTGAGA CTTAATTGGC 540
1138
1139 TGGTGGTATG GTGATATGAT TGACAATAAC TGCTAAGTCA GAGAGGGATA TATTAAGGAG 600
1140
1141 GAGAAGAAAA GCAACAAATC TGGTTTTGAT GTGTTCACTT TGTTATAATT ATTGATTATT 660
1142
1143 TACTGAATAT GAATATTTAT CTTTGTTTTT GAGTCAATAA ATATACCTTT GTAAAGACAG 720
1144
1145 AATTAAAGTA TTAGTATTTC TTTCAAACTG GAGGCATTTC TCCCACTAAC ATATTTCATC 780
1146
1147 AAAACTTATA ATAAGCTTGG TTCCAGAGGA AGAAATGAGG GATAACCAAA AATAGAGACA 840
1148
1149 TTAATAATAG TGTAACGCCC AGTGATAAAT CTCAATAGGC AGTGATGACA GACATGTTTT 900
1150
1151 CCCAAACACA AGGATGCTGT AAGGGCCAAA CAGAAATGAT GGCCCCTCCC CAGCACCTCA 960
1152
1153 TTTTGCCCCT TCCTTCAGCT ATGCCTCTAC TCTCCTTTAG ATACAAGGGA GGTGGATTTT 1020
1154
1155 TCTCTTCTCT GAGATAGCTT GATGGAACCA CAGGAACAAT GAAGTGGGCT CCTGGCTCTT 1080
1156
1157 TTCTCTGTGG CAGATGGGGT GCCATGCCCA CCTTCAGACA AAGGGAAGAT TGAGCTCAAA 1140
1158
1159 AGCTCCCTGA GAAGTGAGAG CCTATGAACA TGGTTGACAC AGAGGGACAG GAATGTATTT 1200
1160
1161 CCAGGGTCAT TCATTCCTGG GAATAGTGAA CTGGGACATG GGGGAAGTCA GTCTCCTCCT 1260
1162
1163 GCCACAGCCA CAGATTAAAA ATAATAATGT TAACTGATCC CTAGGCTAAA ATAATAGTGT 1320
1164
1165 TAACTGATCC CTAAGCTAAG AAAGTTCTTT TGGTAATTCA GGTGATGGCA GCAGGACCCA 1380
1166
1167 TCTTAAGGAT AGACTAGGTT TGCTTAGTTC GAGGTCATAT CTGTTTGCTC TCAGCCATGT 1440
1168
1169 ACTGGAAGAA GTTGCATCAC ACAGCCTCCA GGACTGCCCT CCTCCTCACA GCAATGGATA 1500
1170
1171 ATGCTTCACT AGCCTTTGCA GATAATTTTG GATCAGAGAA AAAACCTTGA GCTGGGCCAA 1560
1172
1173 AAAGGAGGAG CTTCAACCTG TGTGCAAAAT CTGGGAACCT GACAGTATAG GTTGGGGGCC 1620

```
1174
1175  AGGATGAGGA AAAAGGAACG GGAAAGACCT GCCCACCCTT CTGGTAAGGA GGCCCCGTGA    1680
1176
1177  TCAGCTCCAG CCATTTGCAG TCCTGGCTAT CCCAGGAGCT TACATAAAGG GACAATTGGA    1740
1178
1179  GCCTGAGAGG TGACAGTGCT GACACTACAA GGCTCGGAGC TCCGGGCACT CAGACATC      1798
1180
1181  ATG AGT TGG TCC TTG CAC CCC CGG AAT TTA ATT CTC TAC TTC TAT GCT      1846
1182  Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala
1183    1               5                  10                  15
1184
1185  CTT TTA TTT CTC TCT TCA ACA TGT GTA GCA GTAAGTGTGC TCTTCACAAA        1896
1186  Leu Leu Phe Leu Ser Ser Thr Cys Val Ala
1187               20                  25
1188
1189  ACGTTGTTTA AAATGGAAAG CTGGAAAATA AAACAGATAA TAAACTAGTG AAATTTTCGT    1956
1190
1191  ATTTTTTCTC TTTTAG TAT GTT GCT ACC AGA GAC AAC TGC TGC ATC TTA        2005
1192                    Tyr Val Ala Thr Arg Asp Asn Cys Cys Ile Leu
1193                                30                  35
1194
1195  GAT GAA AGA TTC GTAAGTAGTT TTTATGTTTC TCCCTTTGTG TGTGAACTGG          2057
1196  Asp Glu Arg Phe
1197           40
1198
1199  AGAGGGGCAG AGGAATAGAA ATAATTCCCT CATAAATATC ATCTGGCACT TGTAACTTTT    2117
1200
1201  TAAAAACATA GTCTAGGTTT TACCTATTTT TCTTAATAGA TTTTAAGAGT AGCATCTGTC    2177
1202
1203  TACATTTTTA ATCACTGTTA TATTTTCAG GGT AGT TAT TGT CCA ACT ACC TGT      2230
1204                                  Gly Ser Tyr Cys Pro Thr Thr Cys
1205                                              45
1206
1207  GGC ATT GCA GAT TTC CTG TCT ACT TAT CAA ACC AAA GTA GAC AAG GAT      2278
1208  Gly Ile Ala Asp Phe Leu Ser Thr Tyr Gln Thr Lys Val Asp Lys Asp
1209   50                  55                  60                  65
1210
1211  CTA CAG TCT TTG GAA GAC ATC TTA CAT CAA GTT GAA AAC AAA ACA TCA      2326
1212  Leu Gln Ser Leu Glu Asp Ile Leu His Gln Val Glu Asn Lys Thr Ser
1213                   70                  75                  80
1214
1215  GAA GTC AAA CAG CTG ATA AAA GCA ATC CAA CTC ACT TAT AAT CCT GAT      2374
1216  Glu Val Lys Gln Leu Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro Asp
1217               85                  90                  95
1218
1219  GAA TCA TCA AAA CCA  A GTGAGAAAAT AAAGACTACT GACCAAAAAA              2420
1220  Glu Ser Ser Lys Pro
1221          100
1222
1223  TAATAATAAT AATCTGTGAA GTTCTTTTGC TGTTGTTTTA GTTGTTCTAT TTGCTTAAGG    2480
1224
```

```
1225  ATTTTTATGT CTCTGATCCT ATATTACAG  AT ATG ATA GAC GCT GCT ACT TTG      2532
1226                                 Asn Met Ile Asp Ala Ala Thr Leu
1227                                         105                 110
1228
1229  AAG TCC AGG ATA ATG TTA GAA GAA ATT ATG AAA TAT GAA GCA TCG ATT      2580
1230  Lys Ser Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile
1231                  115                 120                 125
1232
1233  TTA ACA CAT GAC TCA AGT ATT  CG  GTAAGGATTT TTGTTTTAAT TTGCTCTGCA    2633
1234  Leu Thr His Asp Ser Ser Ile  Arg
1235              130
1236
1237  AGACTGATTT AGTTTTTATT TAATATTCTA TACTTGAGTG AAAGTAATTT TTAATGTGTT    2693
1238
1239  TTCCCCATTT ATAATATCCC AGTGACATTA TGCCTGATTA TGTTGAGCAT AGTAGAGATA    2753
1240
1241  GAAGTTTTTA GTGCAATATA AATTATACTG GGTTATAATT GCTTATTAAT AATCACATTG    2813
1242
1243  AAGAAAGATG TTCTAGATGT CTTCAAATGC TAGTTTGACC ATATTTATCA AAAATTTTTT    2873
1244
1245  CCCCATCCCC CATTTATCTT ACAACATAAA ATCAATCTCA TAGGAATTTG GGTGTTGAAA    2933
1246
1247  ATAAAATCCT CTTTATAAAA ATGCTGACAA ATTGGTGGTT AAAAAAATTA GCAAGCAGAG    2993
1248
1249  GCATAGTAAG GATTTTGGCT CCTAAAGTAA ATTATATTGA ATGTGGAGCA GGAAGAAACA    3053
1250
1251  TGTCTTGAGA GACTAAGTGT GGCAAATATT GCAAAGCTCA TATTGATCAT TGCAGAATGA    3113
1252
1253  ACCTGCATAG TCTCTTCCCT TCATTTGGAA GTGAATGTCT CTGTTAAAGC TTCTCAGGGA    3173
1254
1255  CTCATAAACT TTCTGAACAT AAGGTCTCAG ATACAGTTTT AATATTTTTC CCCAATTTTT    3233
1256
1257  TTTTCTGAAT TTTTCTCAAA GCAGCTTGAG AAATTGAGAT AAATAGTAGC TAGGGAGAAG    3293
1258
1259  TGGCCCAGGA AAGATTTCTC CTCTTTTTGC TATCAGAGGG CCCTTGTTAT TATTGTTATT    3353
1260
1261  ATTATTACTT GCATTATTAT TGTCCATCAT TGAAGTTGAA GGAGGTTATT GTACAGAAAT    3413
1262
1263  TGCCTAAGAC AAGGTAGAGG GAAAACGTGG ACAAATAGTT TGTCTACCCT TTTTTACTTC    3473
1264
1265  AAAGAAAGAA CGGTTTATGC ATTGTAGACA GTTTTCTATC ATTTTTGGAT ATTTGCAAGC    3533
1266
1267  CACCCTGTAA GTAACTACAA AAGGAGGGTT TTTACTTCCC CCAGTCCATT CCCAAAGCTA    3593
1268
1269  TGTAACCAGA AGCATTAAAG AAGAAAGGGG AAGTATCTGT TGTTTTATTT TACATACAAT    3653
1270
1271  AACGTTCCAG ATCATGTCCC TGTGTAAGTT ATATTTTAGA TTGAAGCTTA TATGTATAGC    3713
1272
1273  CTCAGTAGAT CCACAAGTGA AAGGTATACT CCTTCAGCAC ATGTGAATTA CTGAACTGAG    3773
1274
1275  CTTTTCCTGC TTCTAAAGCA TCAGGGGGTG TTCCTATTAA CCAGTCTCGC CACTCTTGCA    3833
```

```
1276
1277  GGTTGCTATC TGCTGTCCCT TATGCATAAA GTAAAAAGCA AAATGTCAAT GACATTTGCT    3893
1278
1279  TATTGACAAG GACTTTGTTA TTTGTGTTGG GAGTTGAGAC AATATGCCCC ATTCTAAGTA    3953
1280
1281  AAAAGATTCA GGTCCACATT GTATTCCTGT TTTAATTGAT TTTTTGATTT GTTTTTCTTT    4013
1282
1283  TTCAAAAAGT TTATAATTTT AATTCATGTT AATTTAGTAA TATAATTTTA CATTTTCCTC    4073
1284
1285  AAGAATGGAA TAATTTATCA GAAAGCACTT CTTAAGAAAA TACTTAGCAG TTTCCAAAGA    4133
1286
1287  AAATATAAAA TTACTCTTCT GAAAGGAATA CTTATTTTTG TCTTCTTATT TTTGTTATCT    4193
1288
1289  TATGTTTCTG TTTGTAG A TAT TTG CAG GAA ATA TAT AAT TCA AAT AAT CAA      4244
1290                      Tyr Leu Gln Glu Ile Tyr Asn Ser Asn Asn Gln
1291                      135                 140                 145
1292
1293  AAG ATT GTT AAC CTG AAA GAG AAG GTA GCC CAG CTT GAA GCA CAG TGC      4292
1294  Lys Ile Val Asn Leu Lys Glu Lys Val Ala Gln Leu Glu Ala Gln Cys
1295                  150                 155                 160
1296
1297  CAG GAA CCT TGC AAA GAC ACG GTG CAA ATC CAT GAT ATC ACT GGG AAA  G   4341
1298  Gln Glu Pro Cys Lys Asp Thr Val Gln Ile His Asp Ile Thr Gly Lys
1299              165                 170                 175
1300
1301  GTAACTGATG AAGGTTATAT TGGGATTAGG TTCATCAAAG TAAGTAATGT AAAGGAGAAA    4401
1302
1303  GTATGTACTG GAAAGTATAG GAATAGTTTA GAAAGTGGCT ACCCATTAAG TCTAAGAATT    4461
1304
1305  TCAGTTGTCT AGACCTTTCT TGAATAGCTA AAAAAAACAG TTTAAAAGGA ATGCTGATGT    4521
1306
1307  GAAAAGTAAG AAAATTATTC TTGGAAAATG AATAGTTTAC TACATGTTAA AAGCTATTTT    4581
1308
1309  TCAAGGCTGG CACAGTCTTA CCTGCATTTC AAACCACAGT AAAAGTCGAT TCTCCTTCTC    4641
1310
1311  TAG  AT TGT CAA GAC ATT GCC AAT AAG GGA GCT AAA CAG AGC GGG CTT       4688
1312      Asp Cys Gln Asp Ile Ala Asn Lys Gly Ala Lys Gln Ser Gly Leu
1313              180                 185                 190
1314
1315  TAC TTT ATT AAA CCT CTG AAA GCT AAC CAG CAA TTC TTA GTC TAC TGT       4736
1316  Tyr Phe Ile Lys Pro Leu Lys Ala Asn Gln Gln Phe Leu Val Tyr Cys
1317          195                 200                 205
1318
1319  GAA ATC GAT GGG TCT GGA AAT GGA TGG ACT GTG TTT CAG AAG               4778
1320  Glu Ile Asp Gly Ser Gly Asn Gly Trp Thr Val Phe Gln Lys
1321      210                 215                 220
1322
1323  GTAATTTTTT CCCCACCATG TGTATTTAAT AAATTCCTAC ATTGTTTCTG CCATATGGCA    4838
1324
1325  GATACTTTTC TAAGCACCTT GTGAACCGTA GCTCATTTAA TCCTTGCAAT AGCCCTAAGA    4898
1326
```

```
1327  GGAAGGTACT TCTGTTACTC CTATTTACAG AAAAGGAAAC TGAGGCACAC AAGGTTAAAT    4958
1328
1329  AACTTGCCCA AGACCACATA ACTAATAAGC AACAGAGTCA GCATTTGAAC CTAGGCAGTA    5018
1330
1331  TAGTTTCAGA GTTTGTGACT TGACTCTATA TTGTACTGGC ACTGACTTTG TAGATTCATG    5078
1332
1333  GTGGCACATA ATCATAGTAC CACAGTGACA AATAAAAAGA AGGAAACTCT TTTGTCAGGT    5138
1334
1335  AGGTCAAGAC CTGAGGTTTC CCATCACAAG ATGAGGAAGC CCAACACCAC CCCCCACCAC    5198
1336
1337  CCCACCACCA TCACCACCCT TTCACACACC AGAGGATACA CTTGGGCTGC TCCAAGACAA    5258
1338
1339  GGAACCTGTG TTGCATCTGC CACTTGCTGA TACCCACTAG GAATCTTGGC TCCTTTACTT    5318
1340
1341  TCTGTTTACC TCCCACCACT GTTATAACTG TTTCTACAGG GGGCGCTCAG AGGGAATGAA    5378
1342
1343  TGGTGGAAGC ATTAGTTGCC AGACACCGAT TGAGCAATGG GTTCCATCAT AAGTGTAAGA    5438
1344
1345  ATCAGTAATA TCCAGCTAGA GTTCTGAAGT CGTCTAGGTG TCTTTTTAAT ATTACCACTC    5498
1346
1347  ATTTAGAATT TATGATGTGC CAGAAACCCT CTTAAGTATT TCTCTTATAT TCTCTCTCAT    5558
1348
1349  GATCCTTGCA GCAACCCTAA GAAGTAACCA TCATTTTTCC TATTTGATAC ATGAGGAAAC    5618
1350
1351  TGAGGTAGCT TGGCCAAGAT CACTTAGTTG GGAGTTGATA GAACCAGTGC TCTGTATTTT    5678
1352
1353  TGACAAAATG TTGACAGCAT TCTCTTTACA TGCATTGATA GTCTATTTTC TCCTTTTGCT    5738
1354
1355  CTTGCAAATG TGTAATTAG AGA CTT GAT GGC AGT GTA GAT TTC AAG AAA AAC      5790
1356                      Arg Leu Asp Gly Ser Val Asp Phe Lys Lys Asn
1357                              225                 230
1358
1359  TGG ATT CAA TAT AAA GAA GGA TTT GGA CAT CTG TCT CCT ACT GGC ACA      5838
1360  Trp Ile Gln Tyr Lys Glu Gly Phe Gly His Leu Ser Pro Thr Gly Thr
1361      235                 240                 245
1362
1363  ACA GAA TTT TGG CTG GGA AAT GAG AAG ATT CAT TTG ATA AGC ACA CAG      5886
1364  Thr Glu Phe Trp Leu Gly Asn Glu Lys Ile His Leu Ile Ser Thr Gln
1365  250                 255                 260                 265
1366
1367  TCT GCC ATC CCA TAT GCA TTA AGA GTG GAA CTG GAA GAC TGG AAT GGC      5934
1368  Ser Ala Ile Pro Tyr Ala Leu Arg Val Glu Leu Glu Asp Trp Asn Gly
1369                  270                 275                 280
1370
1371  AGA ACC  AG  GTACTGTTTT GAAATGACTT CCAACTTTTT ATTGTAAAGA             5982
1372  Arg Thr  Ser
1373
1374
1375  TTGCCTGGAA TGTGCACTTT CCAACTATCA ATAGACAATG GCAAATGCAG CCTGACAAAT    6042
1376
1377  GCAAACAGCA CATCCAGCCA CCATTTTCTC CAGGAGTCTG TTTGGTTCTT GGGCAATCCA    6102
```

1378
1379 AAAAGGTAAA TTCTATTCAG GATGAATCTA AGTGTATTGG TACAATCTAA TTACCCTGGA 6162
1380
1381 ACCATTCAGA GTAATAGCTA ATTACTGAAC TTTTAATCAG TCCCAGGAAT TGAGCATAAA 6222
1382
1383 ATTATAATTT TATCTAGTCT AAATTACTAT TTCATGAAGC AGGTATTATT ATTAATCCCA 6282
1384
1385 TTTTATAGAT TAACTTGCTC AAAGTCACAT TGCTGATAAG TGGTAGAGGT AGAATTCAGA 6342
1386
1387 CTCAAGTAGT TTAACTTTAG AGCCTGTCCT CTTAACAACT ATCCTGGTTG AAAAGCAAAT 6402
1388
1389 ACAGCCTCTT CAGACTTCTC AGTGCCTTGA TGGCCATTTA TTCTGTCAAA TCATGAGCTA 6462
1390
1391 CCCTAAAAGT AAACCAGCTA GCTCTTTTGA TGATCTAGAG GCTTCTTTTT GCTTGAGATA 6522
1392
1393 TTTGAAGGTT TTAAGCATTG TTACCTAATT AAAATGCAGA AAAATATCCA ACCCTCTTGT 6582
1394
1395 TATGTTTAAG GAATAGTGAA ATATATTGTC TTCAAACACA TGGACTTTTT TTTATTGCTT 6642
1396
1397 GGTTGGTTTT TAATCCAGAA AGTGCTATAG TCAGTAGACC TTCTTCTAGG AAAGGACCTT 6702
1398
1399 CCATTTCCCA GCCACTGGAG ATTAGAAAAT AAGCTAAATA TTTTCTGGAA ATTTCTGTTC 6762
1400
1401 ATTCATTAAG GCCCATCCTT TCCCCCACTC TATAGAAGTG TTGTCCACTT GCACAATTTT 6822
1402
1403 TTCCAGGAAA GAATCTCTCT AACTCCTTCA GCTCACATGC TTTGGACCAC ACAGGGAAGA 6882
1404
1405 CTTTGATTGT GTAATGCCCT CAGAAGCTCT CCTTCTTGCC ACTACCACAC TGATTTGAGG 6942
1406
1407 AAGAAAATCC CTTTAGCACC TAACCCTTCA GGTGCTATGA GTGGCTAATG GAACTGTACC 7002
1408
1409 TCCTTCAAGT TTTGTGCAAT AATTAAGGGT CACTCACTGT CAGATACTTT CTGTGATCTA 7062
1410
1411 TGATAATGTG TGTGCAACAC ATAACATTTC AATAAAAGTA GAAAATATGA AATTAGAGTC 7122
1412
1413 ATCTACACAT CTGGATTTGA TCTTAGAATG AAACAAGCAA AAAAGCATCC AAGTGAGTGC 7182
1414
1415 AATTATTAGT TTTCAGAGAT GCTTCAAAGG CTTCTAGGCC CATCCCGGGA AGTGTTAATG 7242
1416
1417 AGCTGTGGAC TGGTTCACAT ATCTATTGCC TCTTGCCAGA TTTGCAAAAA ACTTCACTCA 7302
1418
1419 ATGAGCAAAT TTCAGCCTTA AGAAACAAAG TCAAAAATTC CAAGGAAGCA TCCTACGAAA 7362
1420
1421 GAGGGAACTT CTGAGATCCC TGAGGAGGGT CAGCATGTGA TGGTTGTATT TCCTTCTTCT 7422
1422
1423 CAG T ACT GCA GAC TAT GCC ATG TTC AAG GTG GGA CCT GAA GCT GAC 7468
1424 Thr Ala Asp Tyr Ala Met Phe Lys Val Gly Pro Glu Ala Asp
1425 285 290 295
1426
1427 AAG TAC CGC CTA ACA TAT GCC TAC TTC GCT GGT GGG GAT GCT GGA GAT 7516
1428 Lys Tyr Arg Leu Thr Tyr Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp

```
1429      300                 305                 310
1430
1431   GCC TTT GAT GGC TTT GAT TTT GGC GAT GAT CCT AGT GAC AAG TTT TTC      7564
1432   Ala Phe Asp Gly Phe Asp Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe
1433   315                 320                 325                 330
1434
1435   ACA TCC CAT AAT GGC ATG CAG TTC AGT ACC TGG GAC AAT GAC AAT GAT      7612
1436   Thr Ser His Asn Gly Met Gln Phe Ser Thr Trp Asp Asn Asp Asn Asp
1437                 335                 340                 345
1438
1439   AAG TTT GAA GGC AAC TGT GCT GAA CAG GAT GGA TCT GGT TGG TGG ATG      7660
1440   Lys Phe Glu Gly Asn Cys Ala Glu Gln Asp Gly Ser Gly Trp Trp Met
1441               350                 355                 360
1442
1443   AAC AAG TGT CAC GCT GGC CAT CTC AAT GGA GTT TAT TAC CAA  G          7703
1444   Asn Lys Cys His Ala Gly His Leu Asn Gly Val Tyr Tyr Gln
1445           365                 370                 375
1446
1447   GTATGTTTTC CTTTCTTAGA TTCCAAGTTA ATGTATAGTG TATACTATTT TCATAAAAAA     7763
1448
1449   TAATAAATAG ATATGAAGAA ATGAAGAATA ATTTATAAAG ATAGTAGGGA TTTTATCATG     7823
1450
1451   TTCTTTATTT CAACTAAGTT CTTTGAAACT GGAAGTGGAT AATACCAAGT TCATGCCTAA     7883
1452
1453   AATTAGCCCT TCTAAAGAAA TCCACCTGCT GCAAAATATC CAGTAGTTTG GCATTATATG     7943
1454
1455   TGAAACTATC ACCATCATAG CTGGCACTGT GGGTTGTGGG ATCTCCTTTA GACATACAAC     8003
1456
1457   ATAAATGATC TGGATGGATT AACATTACTA CATGGATGCT TGTTGACACA TTAACCTGGC     8063
1458
1459   TTCCCATGAG CTTTGTGTCA GATACACGCA GTGAACAGGT GTTTGGAGGA ACAGAATAAA     8123
1460
1461   GAGAAGGCAA GCACTGGTAA GGGCAGGGGT TTGTGAAAGC TTGAGAGAAG AGACCAGTCT     8183
1462
1463   GAGGACAGTA GACACTTATT TTAGGATGGG GGTTGGATGA GGAGGCTATA GTTTGCTATA     8243
1464
1465   AGCTTGGAAT GGTTTGGAAC ACTGGTTTCA CTCACCTACC CAGCAGTTAT GTGTGGGGAA     8303
1466
1467   GCCTTACCGA TGCTAAAGGA TCCATGTTAC AATAATGGCA TTATTTGGAA ATCCCAGTGG     8363
1468
1469   TATTCCATGA ATAAAACCAC TATGAAGATA ATCCCACTCA ACAGACTCTC CGTTGGAGAA     8423
1470
1471   GGACAGCAAC ACCACCCTGG GAAAGCCAAA CAGTCAGACC AGACCTGTTT AGCATCAGTA     8483
1472
1473   GGACTTCCCT ACCATATCTG CTGGGTAGAT GAGTGAAACC AGTGTTCCAA ACCACTCCGG     8543
1474
1475   GCTTGTAGCA AACCATAGTC TCCTCATCTA CCAAGATGAG CAACCTTACC TCCTGATGTC     8603
1476
1477   CTAGCCAATC ACCAACTAGG AAACTTTGCA CAGTTTATTT AAAGTAACAG TTTGATTTTC     8663
1478
1479   ACAATATTTT TAAATTGGAG AAACATAACT TATCTTTGCA CTCACAAACC ACATAATGAG     8723
```

```
1480
1481  AAGAAACTCT AAGGGAAAAT GCTTGATCTG TGTGACCCGG GGCGCCATGC CAGAGCTGTA    8783
1482
1483  GTTCATGCCA GTGTTGTGCT CTGACAAGCC TTTTACAGAA TTACATGAGA TCTGCTTCCC    8843
1484
1485  TAGGACAAGG AGAAGGCAAA TCAACAGAGG CTGCACTTTA AAATGGAGAC ATAAAATAAC    8903
1486
1487  ATGCCAGAAC CATTTCCTAA AGCTCCTCAA TCAACCAACA AAATTGTGCT TTCAAATAAC    8963
1488
1489  CTGAGTTGAC CTCATCAGGA ATTTTGTGGC TCCTTCTCTT CTAACCTGCC TGAAGAAAGA    9023
1490
1491  TGGTCCACAG CAGCTGAGTC CGGGATGGAT AAGCTTAGGG ACAGAGGCCA ATTAGGGAAC    9083
1492
1493  TTTGGGTTTC TAGCCCTACT AGTAGTGAAT AAATTTAAAG TGTGGATGTG ACTATGAGTC    9143
1494
1495  ACAGCACAGA TGTTGTTTAA TAATATGTTT ATTTTATAAA TTGATATTTT AGGAATCTTT    9203
1496
1497  GGAGATATTT TCAGTTAGCA GATAATACTA TAAATTTTAT GTAACTGGCA ATGCACTTCG    9263
1498
1499  TAATAGACAG CTCTTCATAG ACTTGCAGAG GTAAAAAGAT TCCAGAATAA TGATATGTAC    9323
1500
1501  ATCTACGACT TGTTTTAG  GT GGC ACT TAC TCA AAA GCA TCT ACT CCT AAT       9373
1502                     Gly Gly Thr Tyr Ser Lys Ala Ser Thr Pro Asn
1503                                 380                 385
1504
1505  GGT TAT GAT AAT GGC ATT ATT TGG GCC ACT TGG AAA ACC CGG TGG TAT       9421
1506  Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr Arg Trp Tyr
1507          390                 395                 400
1508
1509  TCC ATG AAG AAA ACC ACT ATG AAG ATA ATC CCA TTC AAC AGA CTC ACA       9469
1510  Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn Arg Leu Thr
1511      405                 410                 415
1512
1513  ATT GGA GAA GGA CAG CAA CAC CAC CTG GGG GGA GCC AAA CAG GTC AGA       9517
1514  Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys Gln Val Arg
1515  420                 425                 430                 435
1516
1517  CCA GAG CAC CCT GCG GAA ACA GAA TAT GAC TCA CTT TAC CCT GAG GAT       9565
1518  Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr Pro Glu Asp
1519                  440                 445                 450
1520
1521  GAT TTG TAGAAAATTA ACTGCTAACT TCTATTGACC CACAAAGTTT CAGAAATTCT       9621
1522  Asp Leu
1523
1524
1525  CTGAAAGTTT CTTCCTTTTT TCTCTTACTA TATTTATTGA TTTCAAGTCT TCTATTAAGG    9681
1526
1527  ACATTTAGCC TTCAATGGAA ATTAAAACTC ATTTAGGACT GTATTTCCAA ATTACTGATA    9741
1528
1529  TCAGAGTTAT TTAAAAATTG TTTATTTGAG GAGATAACAT TTCAACTTTG TTCCTAAATA    9801
1530
```

1531 TATAATAATA AAATGATTGA CTTTATTTGC ATTTTTATGA CCACTTGTCA TTTATTTTGT 9861
1532
1533 CTTCGTAAAT TATTTTCATT ATATCAAATA TTTTAGTATG TACTTAATAA AATAGGAGAA 9921
1534
1535 CATTTTAGAG TTTCAAATTC CCAGGTATTT TCCTTGTTTA TTACCCCTAA ATCATTCCTA 9981
1536
1537 TTTAATTCTT CTTTTTAAAT GGAGAAAATT ATGTCTTTTT AATATGGTTT TTGTTTTGTT 10041
1538
1539 ATATATTCAC AGGCTGGAGA CGTTTAAAAG ACCGTTTCAA AAGAGATTTA CTTTTTTAAA 10101
1540
1541 GGACTTTATC TGAACAGAGA GATATAATAT TTTTCCTATT GGACAATGGA CTTGCAAAGC 10161
1542
1543 TTCACTTCAT TTTAAGAGCA AAAGACCCCA TGTTGAAAAC TCCATAACAG TTTTATGCTG 10221
1544
1545 ATGATAATTT ATCTACATGC ATTTCAATAA ACCTTTTGTT TCCTAAGACT AGATACATGG 10281
1546
1547 TACCTTTATT GACCATTAAA AAACCACCAC TTTTTGCCAA TTTACCAATT ACAATTGGGC 10341
1548
1549 AACCATCAGT AGTAATTGAG TCCTCATTTT ATGCTAAATG TTATGCCTAA CTCTTTGGGA 10401
1550
1551 GTTACAAAGG AAATAGCAAT TATGGCTTTT GCCCTCTAGG AGATACAGGA CAAATACAGG 10461
1552
1553 AAAATACAGC AACCCAAACT GACAATACTC TATACAAGAA CATAATCACT AAGCAGGAGT 10521
1554
1555 CACAGCCACA CAACCAAGAT GCATAGTATC CAAAGTGCAG CTG 10564
1556
1557
1558 (2) INFORMATION FOR SEQ ID NO:6:
1559
1560 (i) SEQUENCE CHARACTERISTICS:
1561 (A) LENGTH: 453 amino acids
1562 (B) TYPE: amino acid
1563 (D) TOPOLOGY: linear
1564
1565 (ii) MOLECULE TYPE: protein
1566
1567 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
1568
1569 Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala
1570 1 5 10 15
1571
1572 Leu Leu Phe Leu Ser Ser Thr Cys Val Ala Tyr Val Ala Thr Arg Asp
1573 20 25 30
1574
1575 Asn Cys Cys Ile Leu Asp Glu Arg Phe Gly Ser Tyr Cys Pro Thr Thr
1576 35 40 45
1577
1578 Cys Gly Ile Ala Asp Phe Leu Ser Thr Tyr Gln Thr Lys Val Asp Lys
1579 50 55 60
1580
1581 Asp Leu Gln Ser Leu Glu Asp Ile Leu His Gln Val Glu Asn Lys Thr

```
1582   65                  70                  75                  80
1583
1584   Ser Glu Val Lys Gln Leu Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro
1585                    85                  90                  95
1586
1587   Asp Glu Ser Ser Lys Pro Asn Met Ile Asp Ala Ala Thr Leu Lys Ser
1588               100                 105                 110
1589
1590   Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile Leu Thr
1591           115                 120                 125
1592
1593   His Asp Ser Ser Ile Arg Tyr Leu Gln Glu Ile Tyr Asn Ser Asn Asn
1594       130                 135                 140
1595
1596   Gln Lys Ile Val Asn Leu Lys Glu Lys Val Ala Gln Leu Glu Ala Gln
1597   145                 150                 155                 160
1598
1599   Cys Gln Glu Pro Cys Lys Asp Thr Val Gln Ile His Asp Ile Thr Gly
1600                    165                 170                 175
1601
1602   Lys Asp Cys Gln Asp Ile Ala Asn Lys Gly Ala Lys Gln Ser Gly Leu
1603               180                 185                 190
1604
1605   Tyr Phe Ile Lys Pro Leu Lys Ala Asn Gln Gln Phe Leu Val Tyr Cys
1606           195                 200                 205
1607
1608   Glu Ile Asp Gly Ser Gly Asn Gly Trp Thr Val Phe Gln Lys Arg Leu
1609       210                 215                 220
1610
1611   Asp Gly Ser Val Asp Phe Lys Lys Asn Trp Ile Gln Tyr Lys Glu Gly
1612   225                 230                 235                 240
1613
1614   Phe Gly His Leu Ser Pro Thr Gly Thr Thr Glu Phe Trp Leu Gly Asn
1615                    245                 250                 255
1616
1617   Glu Lys Ile His Leu Ile Ser Thr Gln Ser Ala Ile Pro Tyr Ala Leu
1618               260                 265                 270
1619
1620   Arg Val Glu Leu Glu Asp Trp Asn Gly Arg Thr Ser Thr Ala Asp Tyr
1621           275                 280                 285
1622
1623   Ala Met Phe Lys Val Gly Pro Glu Ala Asp Lys Tyr Arg Leu Thr Tyr
1624       290                 295                 300
1625
1626   Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp Ala Phe Asp Gly Phe Asp
1627   305                 310                 315                 320
1628
1629   Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe Thr Ser His Asn Gly Met
1630                    325                 330                 335
1631
1632   Gln Phe Ser Thr Trp Asp Asn Asp Asn Asp Lys Phe Glu Gly Asn Cys
```

```
1633                340                 345                 350
1634
1635   Ala Glu Gln Asp Gly Ser Gly Trp Trp Met Asn Lys Cys His Ala Gly
1636           355                 360                 365
1637
1638   His Leu Asn Gly Val Tyr Tyr Gln Gly Gly Thr Tyr Ser Lys Ala Ser
1639       370                 375                 380
1640
1641   Thr Pro Asn Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr
1642   385                 390                 395                 400
1643
1644   Arg Trp Tyr Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn
1645                   405                 410                 415
1646
1647   Arg Leu Thr Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys
1648               420                 425                 430
1649
1650   Gln Val Arg Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr
1651           435                 440                 445
1652
1653   Pro Glu Asp Asp Leu
1654       450
1655
1656   (2) INFORMATION FOR SEQ ID NO:7:
1657
1658        (i) SEQUENCE CHARACTERISTICS:
1659             (A) LENGTH: 10807 base pairs
1660             (B) TYPE: nucleic acid
1661             (C) STRANDEDNESS: double
1662             (D) TOPOLOGY: linear
1663
1664
1665      (vii) IMMEDIATE SOURCE:
1666             (B) CLONE: ovine beta-lactoglobulin
1667
1668
1669       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
1670
1671   ACGCGTGTCG ACCTGCAGGT CAACGGATCT CTGTGTCTGT TTTCATGTTA GTACCACACT      60
1672
1673   GTTTTGGTGG CTGTAGCTTT CAGCTACAGT CTGAAGTCAT AAAGCCTGGT ACCTCCAGCT     120
1674
1675   CTGTTCTCTC TCAAGATTGT GTTCTGCTGT TTGGGTCTTT AGTGTCTCCA CACAATTTTT     180
1676
1677   AGAATTGTTT GTTCTAGTTC TGTGAAAAAT GATGCTGGTA TTTTGATAAG GATTGCATTG     240
1678
1679   AATCTGTAAA GCTACAGATA TAGTCATTGG GTAGTACAGT CACTTTAACA ATATTAACTC     300
1680
1681   TTCACATCTG TGAGCATGAT ATATTTTCCC CCTCTATATC ATCTTCAATT CCTCCTATCA     360
1682
1683   GTTTCTTTCA TTGCAGTTTT CTGAGTACAG GTCTTACACC TCCTTGGTTA GAGTCATTCC     420
```

1684
1685 TCAGTATTTT ATTCCTTTGA TACAATTGTG AATGAGGTAA TTTTCTTAGT TTCTCTTTCT 480
1686
1687 GATAGCTCAT TGTTAGTGTA TATATAGAAA AGCAACAGAT TTCTATGTAT TAATTTTGTA 540
1688
1689 TCCTGCAACA GATTTCTATG TATTAATTTT GTATCCTGCT ACTTTACGGA ATTCACTTAT 600
1690
1691 TAGCTTTTTG GTGACATCTT GAGGATTTTC TGAAGAAAAT GGCATGGTAT GGTAGGACAA 660
1692
1693 GGTGTCATGT CATCTGCAAA CAGTGGCAGT TTTCCTTCTT CCCTTCCAAC CTGGATTTCT 720
1694
1695 TTGATTTCTT TCTGTCTGAG TACGACTAGG ATTCCCAATA CTATACCGAA TAAAAGTGGC 780
1696
1697 AAGAGTGGAC ATCCTTGTCT TATTTTTCTG ACCTTAGAGG AAATGCTTTC AGTTTTTCAC 840
1698
1699 CATTAATTAT AATGTTTACT GTGGGCTTGT CATATGTGGC CTTCATTATA TGGAGGTCTA 900
1700
1701 TTCCCTCTAT ACCCACCTTG TTGAGAGTTT TTATCATAAA AGTATGTTGA ATTTTGTCAA 960
1702
1703 AAGTTTTTCC TGCATCTATT GAGATGATTT TTACTCTTCA ATTCATTAAT GATTTTTATT 1020
1704
1705 CTTCATTTTG TTAATGATTT CCATTCTTCA ATTTGTTAAC GTGGTATATC ACATTGATTG 1080
1706
1707 ATTTGTGGAT ACCTTTGTAT CCCTGGGATA AACCTCACTT GATCATGAGC TTTCAATGTA 1140
1708
1709 TTTTTGAATT CACTTTGCTA ATATTCTGTT GGGTATTTTT GCATCTCTAT TCATCAATGA 1200
1710
1711 TATTGGCCTA AGAAAGGTTT TGTCTGGTTT TAGTATCAGG GTGATGCTGG CCTCATAGAG 1260
1712
1713 AGAGTTTAGA AGCATTTCCT CCTCTTTGAT TTTTCGGAAT AGTTTGAGTA GGATAGGTAT 1320
1714
1715 TAACTCTTCT TTAAATGTTT GGGGACTTCC CTGGTGAGCC GGTGGTTGAG AATCCGCCTC 1380
1716
1717 AGGGATGTGG GTTTGATCCC TGGTCAGGGA ACCATTAATA AGATCCCACA TGCTGCAGGC 1440
1718
1719 AACAAGCCCC CAAGCTGCAA CCACTGAGCT GCAACCGCTG CAGTGCCCAC AGGCCACGAC 1500
1720
1721 CAGAGAAAGC CCACATACAG CAGGGAAGAC CCAGCACAAC CGGAAAAAGG AGTTTGGTGG 1560
1722
1723 AATACAGCTG TGAAGCCGTC TGGTCCTGGA CTCCTGCTTG AGGGAATTTT TTAAAAATTA 1620
1724
1725 TTGATTCAAT TTCATTACTG GTAACTGGTC TGTTCATATT TTCTATTTCT TCCGGGTTCA 1680
1726
1727 GTCTTGGGAG ATTGTACATG CCTAGGAATG TGTCCGTTTC TTCTAGGTTG TCCATTTTAT 1740
1728
1729 TGGACATGCA TGGGAGCACA CAGCACCGAC CAGCGAGACT CATGCTGGCT TCCTGGGGCC 1800
1730
1731 AGGCTGGGGC CCCAAGCAGC ATGGCATCCT AGAGTGTGTG AAAGCCCACT GACCCTGCCC 1860
1732
1733 AGCCCCACAA TTTCATTCTG AGAAGTGATT CCTTGCTTCT GCACTTACAG GCCCAGGATC 1920
1734

1735 TGACCTGCTT CTGAGGAGCA GGGGTTTTGG CAGGACGGGG AGATGCTGAG AGCCGACGGG 1980
1736
1737 GGTCCAGGTC CCCTCCCAGG CCCCCCTGTC TGGGGCAGCC CTTGGGAAAG ATTGCCCCAG 2040
1738
1739 TCTCCCTCCT ACAGTGGTCA GTCCCAGCTG CCCCAGGCCA GAGCTGCTTT ATTTCCGTCT 2100
1740
1741 CTCTCTCTGG ATGGTATTCT CTGGAAGCTG AAGGTTCCTG AAGTTATGAA TAGCTTTGCC 2160
1742
1743 CTGAAGGGCA TGGTTTGTGG TCACGGTTCA CAGGAACTTG GGAGACCCTG CAGCTCAGAC 2220
1744
1745 GTCCCGAGAT TGGTGGCACC CAGATTTCCT AAGCTCGCTG GGGAACAGGG CGCTTGTTTC 2280
1746
1747 TCCCTGGCTG ACCTCCCTCC TCCCTGCATC ACCCAGTTCT GAAAGCAGAG CGGTGCTGGG 2340
1748
1749 GTCACAGCCT CTCGCATCTA ACGCCGGTGT CCAAACCACC CGTGCTGGTG TTCGGGGGGC 2400
1750
1751 TACCTATGGG GAAGGGCTTC TCACTGCAGT GGTGCCCCCC GTCCCCTCTG AGATCAGAAG 2460
1752
1753 TCCCAGTCCG GACGTCAAAC AGGCCGAGCT CCCTCCAGAG GCTCCAGGGA GGGATCCTTG 2520
1754
1755 CCCCCCCGCT GCTGCCTCCA GCTCCTGGTG CCGCACCCTT GAGCCTGATC TTGTAGACGC 2580
1756
1757 CTCAGTCTAG TCTCTGCCTC CGTGTTCACA CGCCTTCTCC CCATGTCCCC TCCGTGTCCC 2640
1758
1759 CGTTTTCTCT CACAAGGACA CCGGACATTA GATTAGCCCC TGTTCCAGCC TCACCTGAAC 2700
1760
1761 AGCTCACATC TGTAAAGACC TAGATTCCAA ACAAGATTCC AACCTGAAGT TCCCGGTGGA 2760
1762
1763 TGTGAGTTCT GGGGCGACAT CCTTCAACCC CATCACAGCT TGCAGTTCAT CGCAAAACAT 2820
1764
1765 GGAACCTGGG GTTTATCGTA AAACCCAGGT TCTTCATGAA ACACTGAGCT TCGAGGCTTG 2880
1766
1767 TTGCAAGAAT TAAAGGTGCT AATACAGATC AGGGCAAGGA CTGAAGCTGG CTAAGCCTCC 2940
1768
1769 TCTTTCCATC ACAGGAAAGG GGGGCCTGGG GGCGGCTGGA GGTCTGCTCC CGTGAGTGAG 3000
1770
1771 CTCTTTCCTG CTACAGTCAC CAACAGTCTC TCTGGGAAGG AAACCAGAGG CCAGAGAGCA 3060
1772
1773 AGCCGGAGCT AGTTTAGGAG ACCCCTGAAC CTCCACCCAA GATGCTGACC AGCCAGCGGG 3120
1774
1775 CCCCCTGGAA AGACCCTACA GTTCAGGGGG GAAGAGGGGC TGACCCGCCA GGTCCCTGCT 3180
1776
1777 ATCAGGAGAC ATCCCCGCTA TCAGGAGATT CCCCCACCTT GCTCCCGTTC CCCTATCCCA 3240
1778
1779 ATACGCCCAC CCCACCCCTG TGATGAGCAG TTTAGTCACT TAGAATGTCA ACTGAAGGCT 3300
1780
1781 TTTGCATCCC CTTTGCCAGA GGCACAAGGC ACCCACAGCC TGCTGGGTAC CGACGCCCAT 3360
1782
1783 GTGGATTCAG CCAGGAGGCC TGTCCTGCAC CCTCCCTGCT CGGGCCCCCT CTGTGCTCAG 3420
1784
1785 CAACACACCC AGCACCAGCA TTCCCGCTGC TCCTGAGGTC TGCAGGCAGC TCGCTGTAGC 3480

1786
1787 CTGAGCGGTG TGGAGGGAAG TGTCCTGGGA GATTTAAAAT GTGAGAGGCG GGAGGTGGGA 3540
1788
1789 GGTTGGGCCC TGTGGGCCTG CCCATCCCAC GTGCCTGCAT TAGCCCCAGT GCTGCTCAGC 3600
1790
1791 CGTGCCCCCG CCGCAGGGGT CAGGTCACTT TCCCGTCCTG GGGTTATTAT GACTCTTGTC 3660
1792
1793 ATTGCCATTG CCATTTTTGC TACCCTAACT GGGCAGCAGG TGCTTGCAGA GCCCTCGATA 3720
1794
1795 CCGACCAGGT CCTCCCTCGG AGCTCGACCT GAACCCCATG TCACCCTTGC CCCAGCCTGC 3780
1796
1797 AGAGGGTGGG TGACTGCAGA GATCCCTTCA CCCAAGGCCA CGGTCACATG GTTTGGAGGA 3840
1798
1799 GCTGGTGCCC AAGGCAGAGG CCACCCTCCA GGACACACCT GTCCCCAGTG CTGGCTCTGA 3900
1800
1801 CCTGTCCTTG TCTAAGAGGC TGACCCCGGA AGTGTTCCTG GCACTGGCAG CCAGCCTGGA 3960
1802
1803 CCCAGAGTCC AGACACCCAC CTGTGCCCCC GCTTCTGGGG TCTACCAGGA ACCGTCTAGG 4020
1804
1805 CCCAGAGGGG ACTTCCTGCT TGGCCTTGGA TGGAAGAAGG CCTCCTATTG TCCTCGTAGA 4080
1806
1807 GGAAGCCACC CCGGGGCCTG AGGATGAGCC AAGTGGGATT CCGGGAACCG CGTGGCTGGG 4140
1808
1809 GGCCCAGCCC GGGCTGGCTG GCCTGCATGC CTCCTGTATA AGGCCCCAAG CCTGCTGTCT 4200
1810
1811 CAGCCCTCCA CTCCCTGCAG AGCTCAGAAG CACGACCCCA GGGATATCCC TGCAGCCATG 4260
1812
1813 AAGTGCCTCC TGCTTGCCCT GGGCCTGGCC CTCGCCTGTG GCGTCCAGGC CATCATCGTC 4320
1814
1815 ACCCAGACCA TGAAAGGCCT GGACATCCAG AAGGTTCGAG GGTTGGCCGG GTGGGTGAGT 4380
1816
1817 TGCAGGGCGG GCAGGGGAGC TGGGCCTCAG AGAGCCAAGA GAGGCTGTGA CGTTGGGTTC 4440
1818
1819 CCATCAGTCA GCTAGGGCCA CCTGACAAAT CCCCGCTGGG GCAGCTTCAA CCAGGCGTTC 4500
1820
1821 ACTGTCTTGC ATTCTGGAGG CTGGAAGCCC AAGATCCAGG TGTTGGCAGG GCTGGCTTCT 4560
1822
1823 CCTGCGGCCG CTCTCTGGGG AGCAGACGGC CGTCTTCTCC AGTCCTCTGC GCGCCCTGAT 4620
1824
1825 TTCCTCTTCC TGTGAGGCCA CCAGGCCTGC TGGAAACACG CCTGCCTGCG CAGCTTCACA 4680
1826
1827 CGACCTTTGT CATCTCTTTA AAGGCCATGT CTCCAGAGTC ATGTGTTGAA GTTCTGGGGG 4740
1828
1829 TTAGTGGGAC ACAGTTCAGC CCCTAAAAGA GTCTCTCTGC CCCTCAAATT TTCCCCACCT 4800
1830
1831 CCAGCCATGT CTCCCCAAGA TCCAAATGTT GCTACATGTG GGGGGGCTCA TCTGGGTCCC 4860
1832
1833 TCTTTGGGTT CAGTGTGAGT CTGGGGAGAG CATTCCCCAG GGTGCAGAGT TGGGGGGAGT 4920
1834
1835 ATCTCAGGGC TGCCCAGGCC GGGGTGGGAC AGAGAGCCCA CTGTGGGGCT GGGGGCCCCT 4980
1836

1837 TCCCACCCCC AGAGTGCAAC TCAAGGTCCC TCTCCAGGTG GCGGGGACTT GGCACTCCTT 5040
1838
1839 GGCTATGGCG GCCAGCGACA TCTCCCTGCT GGATGCCCAG AGTGCCCCCC TGAGAGTGTA 5100
1840
1841 CGTGGAGGAG CTGAAGCCCA CCCCCGAGGG CAACCTGGAG ATCCTGCTGC AGAAATGGTG 5160
1842
1843 GGCGTCTCTC CCCAACATGG AACCCCCACT CCCCAGGGCT GTGGACCCCC CGGGGGGTGG 5220
1844
1845 GGTGCAGGAG GGACCAGGGC CCCAGGGCTG GGGAAGAGGG CTCAGAGTTT ACTGGTACCC 5280
1846
1847 GGCGCTCCAC CCAAGGCTGC CCACCCAGGG CTTTTTTTTT TTTTAAACTT TTATTAATTT 5340
1848
1849 GATGCTTCAG AACATCATCA AACAAATGAA CATAAAACAT TCATTTTTGT TTACTTGGAA 5400
1850
1851 GGGGAGATAA AATCCTCTGA AGTGGAAATG CATAGCAAAG ATACATACAA TGAGGCAGGT 5460
1852
1853 ATTCTGAATT CCCTGTTAGT CTGAGGATTA CAAGTGTATT TGAGCAACAG AGAGACATTT 5520
1854
1855 TCATCATTTC TAGTCTGAAC ACCTCAGTAT CTAAAATGAA CAAGAAGTCC TGGAAACGAA 5580
1856
1857 GCAGTGTGGG GATAGGCCCG TGTGAAGGCT GCTGGGAGGC AGCAGACCTG GGTCTTCGGG 5640
1858
1859 CTCAAGCAGT TCCCGCTACC AGCCCTGTCC ACCTCAGACG GGGGTCAGGG TGCAGGAGAG 5700
1860
1861 AGCTGGATGG GTGTGGGGGC AGAGATGGGG ACCTGAACCC CAGGGCTGCC TTTTGGGGGT 5760
1862
1863 GCCTGTGGTC AAGGCTCTCC CTGACCTTTT CTCTCTGGCT TCATCTGACT TCTCCTGGCC 5820
1864
1865 CATCCACCCG GTCCCCTGTG GCCTGAGGTG ACAGTGAGTG CGCCGAGGCT AGTTGGCCAG 5880
1866
1867 CTGGCTCCTA TGCCCATGCC ACCCCCCTCC AGCCCTCCTG GGCCAGCTTC TGCCCCTGGC 5940
1868
1869 CCTCAGTTCA TCCTGATGAA AATGGTCCAT GCCAATGGCT CAGAAAGCAG CTGTCTTTCA 6000
1870
1871 GGGAGAACGG CGAGTGTGCT CAGAAGAAGA TTATTGCAGA AAAAACCAAG ATCCCTGCGG 6060
1872
1873 TGTTCAAGAT CGATGGTGAG TCCGGGTCCC TGGGGGACAC CCACCACCCC CGCCCCCGGG 6120
1874
1875 GACTGTGGAC AGGTTCAGGG GGCTGGCGTC GGGCCCTGGG ATGCTAAGGG ACTGGTGGTG 6180
1876
1877 ATGAAGACAC TGCCTTGACA CCTGCTTCAC TTGCCTCCCC TGCCACCTGC CCGGGGCCTT 6240
1878
1879 GGGGCGGTGG CCATGGGCAG GTCCCGGCTG GCGGGCTAAC CCACCAGGGT GACACCCGAG 6300
1880
1881 CTCTCTTTGC TGGGGGGCGG GCGGTGCTCT GGGCCCTCAG GCTGAGCTCA GGAGGTACCT 6360
1882
1883 GTGCCCTCCC AGGGGTAACC GAGAGCCGTT GCCCACTCCA GGGGCCCAGG TGCCCCACGA 6420
1884
1885 CCCCAGCCCG CTCCACAGCT CCTTCATCTC CTGGAGACAA ACTCTGTCCG CCCTCGCTCA 6480
1886
1887 TTCACTTGTT CGTCCTAAAT CCGAGATGAT AAAGCTTCGA GGGGGGGTTG GGGTTCCATC 6540

1888
1889 AGGGCTGCCC TTCCGCCGGG CAGCCTGGGC CACATCTGCC CTTGGCCCCC TCAGGACTCA 6600
1890
1891 CTCTGACTGG AGGCCCTGCA CTGACTGACG CCAGGGTGCC CAGCCCAGGG TCTCTGGCGC 6660
1892
1893 CATCCAGCTG CACTGGGTTT GGGTGCTGGT CCTGCCCCCA AGCTGCCCGG ACACCACAGG 6720
1894
1895 CAGCCGGGGC TGCCCACTGG CCTCGGTCAG GGTGAGCCCC AGCTGCCCCC GCTCAGGGCT 6780
1896
1897 TGCCCCGACA ATGACCCCAT CCTCAGGACG CACCCCCCTT CCCTTGCTGG GCAGTGTCCA 6840
1898
1899 GCCCCACCCG AGATCGGGGG AAGCCCTATT TCTTGACAAC TCCAGTCCCT GGGGGAGGGG 6900
1900
1901 GCCTCAGACT GAGTGGTGAG TGTTCCCAAG TCCAGGAGGT GGTGGAGGGT CCTGGCGGAT 6960
1902
1903 CCAGAGTTGA CAGTGAGGGC TTCCTGGGCC CCATGCGCCT GGCAGTGGCA GCAGGGAAGA 7020
1904
1905 GGAAGCACCA TTTCAGGGGT GGGGGATGCC AGAGGCGCTC CCCACCCCGT CTTCGCCGGG 7080
1906
1907 TGGTGACCCC GGGGGAGCCC CGCTGGTCGT GGAGGGTGCT GGGGGCTGAC TAGCAACCCC 7140
1908
1909 TCCCCCCCCG TTGGAACTCA CTTTTCTCCC GTCTTGACCG CGTCCAGCCT TGAATGAGAA 7200
1910
1911 CAAAGTCCTT GTGCTGGACA CCGACTACAA AAAGTACCTG CTCTTCTGCA TGGAAAACAG 7260
1912
1913 TGCTGAGCCC GAGCAAAGCC TGGCCTGCCA GTGCCTGGGT GGGTGCCAAC CCTGGCTGCC 7320
1914
1915 CAGGGAGACC AGCTGCGTGG TCCTTGCTGC AACAGGGGGT GGGGGGTGGG AGCTTGATCC 7380
1916
1917 CCAGGAGGAG GAGGGGTGGG GGGTCCCTGA GTCCCGCCAG GAGAGAGTGG TCGCATACCG 7440
1918
1919 GGAGCCAGTC TGCTGTGGGC CTGTGGGTGG CTGGGGACGG GGGCCAGACA CACAGGCCGG 7500
1920
1921 GAGACGGGTG GGCTGCAGAA CTGTGACTGG TGTGACCGTC GCGATGGGGC CGGTGGTCAC 7560
1922
1923 TGAATCTAAC AGCCTTTGTT ACCGGGGAGT TTCAATTATT TCCCAAAATA AGAACTCAGG 7620
1924
1925 TACAAAGCCA TCTTTCAACT ATCACATCCT GAAAACAAAT GGCAGGTGAC ATTTTCTGTG 7680
1926
1927 CCGTAGCAGT CCCACTGGGC ATTTTCAGGG CCCCTGTGCC AGGGGGGCGC GGGCATCGGC 7740
1928
1929 GAGTGGAGGC TCCTGGCTGT GTCAGCCGGC CCAGGGGGAG GAAGGGACCC GGACAGCCAG 7800
1930
1931 AGGTGGGGGG CAGGCTTTCC CCCTGTGACC TGCAGACCCA CTGCACTGCC CTGGGAGGAA 7860
1932
1933 GGGAGGGGAA CTAGGCCAAG GGGGAAGGGC AGGTGCTCTG GAGGGCAAGG GCAGACCTGC 7920
1934
1935 AGACCACCCT GGGGAGCAGG GACTGACCCC CGTCCCTGCC CCATAGTCAG GACCCCGGAG 7980
1936
1937 GTGGACAACG AGGCCCTGGA GAAATTCGAC AAAGCCCTCA AGGCCCTGCC CATGCACATC 8040
1938

```
1939  CGGCTTGCCT TCAACCCGAC CCAGCTGGAG GGTGAGCACC CAGGCCCCGC CCTTCCCCAG    8100
1940
1941  GGCAGGAGCC ACCCGGCCCC GGGACGACCT CCTCCCATGG TGACCCCCAG CTCCCCAGGC    8160
1942
1943  CTCCCAGGAG GAAGGGGTGG GGTGCAGCAC CCCGTGGGGG CCCCCTCCCC ACCCCCTGCC    8220
1944
1945  AGGCCTCTCT TCCCGAGGTG TCCAGTCCCA TCCTGACCCC CCCATGACTC TCCCTCCCCC    8280
1946
1947  ACAGGGCAGT GCCACGTCTA GGTGAGCCCC TGCCGGTGCC TCTGGGGTAA GCTGCCTGCC    8340
1948
1949  CTGCCCCACG TCCTGGGCAC ACACATGGGG TAGGGGGTCT TGGTGGGGCC TGGGACCCCA    8400
1950
1951  CATCAGGCCC TGGGGTCCCC CCTGTGAGAA TGGCTGGAAG CTGGGGTCCC TCCTGGCGAC    8460
1952
1953  TGCAGAGCTG GCTGGCCGCG TGCCACTCTT GTGGGTGACC TGTGTCCTGG CCTCACACAC    8520
1954
1955  TGACCTCCTC CAGCTCCTTC CAGCAGAGCT AAGGCTAAGT GAGCCAGAAT GGTACCTAAG    8580
1956
1957  GGGAGGCTAG CGGTCCTTCT CCCGAGGAGG GGCTGTCCTG GAACCACCAG CCATGGAGAG    8640
1958
1959  GCTGGCAAGG GTCTGGCAGG TGCCCCAGGA ATCACAGGGG GGCCCCATGT CCATTTCAGG    8700
1960
1961  GCCCGGGAGC CTTGGACTCC TCTGGGGACA GACGACGTCA CCACCGCCCC CCCCCCATCA    8760
1962
1963  GGGGGACTAG AAGGGACCAG GACTGCAGTC ACCCTTCCTG GGACCCAGGC CCCTCCAGGC    8820
1964
1965  CCCTCCTGGG GCTCCTGCTC TGGGCAGCTT CTCCTTCACC AATAAAGGCA TAAACCTGTG    8880
1966
1967  CTCTCCCTTC TGAGTCTTTG CTGGACGACG GGCAGGGGGT GGAGAAGTGG TGGGGAGGGA    8940
1968
1969  GTCTGGCTCA GAGGATGACA GCGGGGCTGG GATCCAGGGC GTCTGCATCA CAGTCTTGTG    9000
1970
1971  ACAACTGGGG GCCCACACAC ATCACTGCGG CTCTTTGAAA CTTTCAGGAA CCAGGGAGGG    9060
1972
1973  ACTCGGCAGA GACATCTGCC AGTTCACTTG GAGTGTTCAG TCAACACCCA AACTCGACAA    9120
1974
1975  AGGACAGAAA GTGGAAAATG GCTGTCTCTT AGTCTAATAA ATATTGATAT GAAACTCAAG    9180
1976
1977  TTGCTCATGG ATCAATATGC CTTTATGATC CAGCCAGCCA CTACTGTCGT ATCAACTCAT    9240
1978
1979  GTACCCAAAC GCACTGATCT GTCTGGCTAA TGATGAGAGA TTCCCAGTAG AGAGCTGGCA    9300
1980
1981  AGAGGTCACA GTGAGAACTG TCTGCACACA CAGCAGAGTC CACCAGTCAT CCTAAGGAGA    9360
1982
1983  TCAGTCCTGG TGTTCATTGG AGGACTGATG TTGAAGCTGA AACTCCAATG CTTTGGCCAC    9420
1984
1985  CTGATGTGAA GAGCTGACTC ATTTGAAAAG ACCCTGATGC TGGGAAAGAT TGAGGGCAGG    9480
1986
1987  AGGAGAAGGG GACGACAGAG GATGAGATGG TTGGATGGCA TCACCAACAC AATGGACATG    9540
1988
1989  GGTTTGGGTG GACTCCAGGA GTTGGTGATG GACAGGGAGG CCTGGCGTGC TACGGAAGCG    9600
```

GTTTATGGGG TCACAAAGAC TGAGTGACTG AACTGAGCTG AACTGAATGG AAATGAGGTA 9660

TACAGCAAAG TGGGGATTTT TTAGATAATA AGAATATACA CATAACATAG TGTATACTCA 9720

TATTTTTATG CATACCTGAA TGCTCAGTCA CTCAGTCGTA TCTGACTCTG TGACCTATGG 9780

ACCGTAGCCT TCCAGGTTTC TTCTGTCCAC AGAATTCTCC AAGGCAAGAA TACTGGAGTG 9840

GGTAGCCATT TCCTCCTCCA GGGGATCCTC CCGACCCAGG GATTGAACCG GCATCTCCTG 9900

TATTGGCAGG TGGATTCTTT ACCACTGTGC CACCAGGGAA GCCCGTGTTA CTCTCTATGT 9960

CCCACTTAAT TACCAAAGCT GCTCCAAGAA AAAGCCCCTG TGCCCTCTGA GCTTCCCGGC 10020

CTGCAGAGGG TGGTGGGGGT AGACTGTGAC CTGGGAACAC CCTCCCGCTT CAGGACTCCC 10080

GGGCCACGTG ACCCACAGTC CTGCAGACAG CCGGGTAGCT CTGCTCTTCA AGGCTCATTA 10140

TCTTTAAAAA AAACTGAGGT CTATTTTGTG ACTTCGCTGC CGTAACTTCT GAACATCCAG 10200

TGCGATGGAC AGGACCTCCT CCCCAGGCCT CAGGGGCTTC AGGGAGCCAG CCTTCACCTA 10260

TGAGTCACCA GACACTCGGG GGTGGCCCCG CCTTCAGGGT GCTCACAGTC TTCCCATCGT 10320

CCTGATCAAA GAGCAAGACC AATGACTTCT TAGGAGCAAG CAGACACCCA CAGGACACTG 10380

AGGTTCACCA GAGCTGAGCT GTCCTTTTGA ACCTAAAGAC ACACAGCTCT CGAAGGTTTT 10440

CTCTTTAATC TGGATTTAAG GCCTACTTGC CCCTCAAGAG GGAAGACAGT CCTGCATGTC 10500

CCCAGGACAG CCACTCGGTG GCATCCGAGG CCACTTAGTA TTATCTGACC GCACCCTGGA 10560

ATTAATCGGT CCAAACTGGA CAAAAACCTT GGTGGGAAGT TTCATCCCAG AGGCCTCAAC 10620

CATCCTGCTT TGACCACCCT GCATCTTTTT TTCTTTTATG TGTATGCATG TATATATATA 10680

TATATATTTT TTTTTTTTTC ATTTTTTGGC TGTGCTGGCT GTTCGTTGCA GTTCGGTGCG 10740

CAGGCTTCTC TCTAGTTTCT CTCTAGTCTT CTCTTATCAC AGAGCAGTCT CTAGACGATC 10800

GACGCGT 10807

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

2041
2042
2043
2044 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
2045
2046 AATTCCGATC GACGCGTCGA CGATATACTC TAGACGATCG ACGCGTA 47
2047
2048 (2) INFORMATION FOR SEQ ID NO:9:
2049
2050 (i) SEQUENCE CHARACTERISTICS:
2051 (A) LENGTH: 24 base pairs
2052 (B) TYPE: nucleic acid
2053 (C) STRANDEDNESS: single
2054 (D) TOPOLOGY: linear
2055
2056
2057 (vii) IMMEDIATE SOURCE:
2058 (B) CLONE: BLGAMP3
2059
2060
2061 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
2062
2063 TGGATCCCCT GCCGGTGCCT CTGG 24
2064
2065
2066 (2) INFORMATION FOR SEQ ID NO:10:
2067
2068 (i) SEQUENCE CHARACTERISTICS:
2069 (A) LENGTH: 24 base pairs
2070 (B) TYPE: nucleic acid
2071 (C) STRANDEDNESS: single
2072 (D) TOPOLOGY: linear
2073
2074
2075 (vii) IMMEDIATE SOURCE:
2076 (B) CLONE: BLGAMP4
2077
2078
2079 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
2080
2081 AACGCGTCAT CCTCTGTGAG CCAG 24
2082
2083 (2) INFORMATION FOR SEQ ID NO:11:
2084
2085 (i) SEQUENCE CHARACTERISTICS:
2086 (A) LENGTH: 10 base pairs
2087 (B) TYPE: nucleic acid
2088 (C) STRANDEDNESS: single
2089 (D) TOPOLOGY: linear
2090
2091

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6839

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTACGTAGT 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6632

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGACGCGGAT CCTACGTACC TGCAGCCATG TTTTCCATGA GG 42

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6627

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGGCTTCGG CAAGCTTCAG G 21

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6521

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCCAAAGACT TACTTCCCTC TAGA 24

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6520

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCATGAACGT CGCGTGGTGG TTGTGCTACC 30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6519

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCACGCGAC GTTCATGCTC TAAAACCGTT 30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

2194
2195 (vii) IMMEDIATE SOURCE:
2196 (B) CLONE: ZC6518
2197
2198
2199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
2200
2201 GCTGCGGGAT CCTACGTACT AGGGGGACAG GGAAGG 36
2202
2203 (2) INFORMATION FOR SEQ ID NO:18:
2204
2205 (i) SEQUENCE CHARACTERISTICS:
2206 (A) LENGTH: 45 base pairs
2207 (B) TYPE: nucleic acid
2208 (C) STRANDEDNESS: single
2209 (D) TOPOLOGY: linear
2210
2211
2212 (vii) IMMEDIATE SOURCE:
2213 (B) CLONE: ZC6629
2214
2215
2216 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
2217
2218 CGACGCGAAT TCTACGTACC TGCAGCCATG AAAAGGATGG TTTCT 45
2219
2220 (2) INFORMATION FOR SEQ ID NO:19:
2221
2222 (i) SEQUENCE CHARACTERISTICS:
2223 (A) LENGTH: 45 base pairs
2224 (B) TYPE: nucleic acid
2225 (C) STRANDEDNESS: single
2226 (D) TOPOLOGY: linear
2227
2228
2229 (vii) IMMEDIATE SOURCE:
2230 (B) CLONE: ZC6630
2231
2232
2233 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
2234
2235 CGACGCGAAT TCTACGTACC TGCAGCCATG AAACATCTAT TATTG 45
2236
2237
2238 (2) INFORMATION FOR SEQ ID NO:20:
2239
2240 (i) SEQUENCE CHARACTERISTICS:
2241 (A) LENGTH: 21 base pairs
2242 (B) TYPE: nucleic acid
2243 (C) STRANDEDNESS: single
2244 (D) TOPOLOGY: linear

2245
2246
2247 (vii) IMMEDIATE SOURCE:
2248 (B) CLONE: ZC6625
2249
2250
2251 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
2252
2253 GTGAGATTTT CAGATCTTGT C 21
2254
2255 (2) INFORMATION FOR SEQ ID NO:21:
2256
2257 (i) SEQUENCE CHARACTERISTICS:
2258 (A) LENGTH: 21 base pairs
2259 (B) TYPE: nucleic acid
2260 (C) STRANDEDNESS: single
2261 (D) TOPOLOGY: linear
2262
2263
2264 (vii) IMMEDIATE SOURCE:
2265 (B) CLONE: ZC6626
2266
2267
2268 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
2269
2270 AAGAATTACT GTGGCCTACC A 21
2271
2272 (2) INFORMATION FOR SEQ ID NO:22:
2273
2274 (i) SEQUENCE CHARACTERISTICS:
2275 (A) LENGTH: 33 base pairs
2276 (B) TYPE: nucleic acid
2277 (C) STRANDEDNESS: single
2278 (D) TOPOLOGY: linear
2279
2280
2281
2282 (vii) IMMEDIATE SOURCE:
2283 (B) CLONE: ZC6624
2284
2285
2286 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
2287
2288 GCTGCGGAAT TCTACGTACT ATTGCTGTGG GAA 33
2289
2290 (2) INFORMATION FOR SEQ ID NO:23:
2291
2292 (i) SEQUENCE CHARACTERISTICS:
2293 (A) LENGTH: 45 base pairs
2294 (B) TYPE: nucleic acid
2295 (C) STRANDEDNESS: single

2296 (D) TOPOLOGY: linear
2297
2298
2299 (vii) IMMEDIATE SOURCE:
2300 (B) CLONE: ZC6514
2301
2302
2303 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
2304
2305 CGACGCGGAT CCTACGTACC TGCAGCCATG AGTTGGTCCT TGCAC 45
2306
2307 (2) INFORMATION FOR SEQ ID NO:24:
2308
2309 (i) SEQUENCE CHARACTERISTICS:
2310 (A) LENGTH: 21 base pairs
2311 (B) TYPE: nucleic acid
2312 (C) STRANDEDNESS: single
2313 (D) TOPOLOGY: linear
2314
2315
2316 (vii) IMMEDIATE SOURCE:
2317 (B) CLONE: zc6517
2318
2319
2320 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
2321
2322 GTCTCTGGTA GCAACATACT A 21
2323
2324
2325 (2) INFORMATION FOR SEQ ID NO:25:
2326
2327 (i) SEQUENCE CHARACTERISTICS:
2328 (A) LENGTH: 22 base pairs
2329 (B) TYPE: nucleic acid
2330 (C) STRANDEDNESS: single
2331 (D) TOPOLOGY: linear
2332
2333
2334 (vii) IMMEDIATE SOURCE:
2335 (B) CLONE: zc6516
2336
2337
2338 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
2339
2340 GGGTTTCTAG CCCTACTAGT AG 22
2341
2342 (2) INFORMATION FOR SEQ ID NO:26:
2343
2344 (i) SEQUENCE CHARACTERISTICS:
2345 (A) LENGTH: 22 base pairs
2346 (B) TYPE: nucleic acid

2347 (C) STRANDEDNESS: single
2348 (D) TOPOLOGY: linear
2349
2350
2351 (vii) IMMEDIATE SOURCE:
2352 (B) CLONE: zc6515
2353
2354
2355 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
2356
2357 GGGTTTCTAG CCCTACTAGT AG 22
2358
2359
2360 (2) INFORMATION FOR SEQ ID NO:27:
2361
2362 (i) SEQUENCE CHARACTERISTICS:
2363 (A) LENGTH: 47 base pairs
2364 (B) TYPE: nucleic acid
2365 (C) STRANDEDNESS: single
2366 (D) TOPOLOGY: linear
2367
2368
2369
2370 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
2371
2372 AAGCTACGCG TCGATCGTCT AGAGTATATC GTCGACGCGT CGATCGG 47
2373
2374

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/206,176*

DATE: 03/23/94
TIME: 15:05:24

INPUT SET: S2548.raw

Line	Error	Original Text

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION *US/08/206,176*

INPUT SET: S2548.raw

APPLICATION NUMBER
FILING DATE
PRIOR APPLICATION DATA

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/08/206,176*

DATE: 03/23/94
TIME: 15:05:25

INPUT SET: S2548.raw

Line	Original Text	Corrected Text